

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:19:27 ; Search time 213.041 Seconds  
(without alignments)  
294.149 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_225\_296

Perfect score: 370

Sequence: 1 KEKTPPEKPEVTKANLIY.....GKYTADLEGGVTINIRFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 segs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA.Main.\*

```
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370	100.0	72	23	US-09-808-212A-10
2	370	100.0	305	7	US-08-325-278-1

3	370	100.0	305	7	US-08-325-278A-1	Sequence 1, Appli
4	370	100.0	305	7	US-08-325-278B-1	Sequence 1, Appli
5	370	100.0	434	7	US-08-325-278-3	Sequence 3, Appli
6	370	100.0	434	7	US-08-325-278A-3	Sequence 3, Appli
7	370	100.0	434	7	US-08-325-278B-3	Sequence 3, Appli
8	370	100.0	467	25	US-09-980-469-12	Sequence 12, Appli
9	370	100.0	719	22	US-09-791-537-10210	Sequence 10210, A
10	370	100.0	719	22	US-09-791-537-96101	Sequence 96101, A
11	339	91.6	72	23	US-09-808-212A-8	Sequence 8, Appli
12	329	88.9	75	23	US-09-808-212A-18	Sequence 18, Appli
13	329	88.9	291	8	US-08-446-137A-4	Sequence 4, Appli
14	329	88.9	992	22	US-09-791-537-88366	Sequence 88366, A
15	329	88.9	1027	7	US-08-331-637-2	Sequence 2, Appli
16	329	88.9	1027	8	US-08-446-137A-2	Sequence 2, Appli
17	329	88.9	1027	15	US-09-187-295-2	Sequence 2, Appli
18	322	87.0	72	23	US-09-808-212A-6	Sequence 6, Appli
19	322	87.0	72	23	US-09-889-182A-4	Sequence 4, Appli
20	315.5	85.3	71	23	US-09-808-212A-14	Sequence 14, Appli
21	306	82.7	74	23	US-09-808-212A-16	Sequence 16, Appli
22	304	82.2	82	23	US-09-808-212A-2	Sequence 2, Appli
23	269	72.7	76	23	US-09-808-212A-4	Sequence 4, Appli
24	269	72.7	78	22	US-09-791-537-22553	Sequence 22553, A
25	254	68.6	182	19	US-09-509-031-2	Sequence 2, Appli
26	254	68.6	182	23	US-09-820-048A-2	Sequence 2, Appli
27	251.5	68.0	71	23	US-09-808-212A-12	Sequence 12, Appli
28	246	66.5	178	19	US-09-509-031-13	Sequence 13, Appli
29	246	66.5	178	23	US-09-820-048A-13	Sequence 13, Appli
30	246	66.5	198	19	US-09-509-031-8	Sequence 8, Appli
31	246	66.5	198	23	US-09-820-048A-8	Sequence 8, Appli
32	246	66.5	342	19	US-09-509-031-6	Sequence 6, Appli
33	246	66.5	342	23	US-09-820-048A-6	Sequence 6, Appli
34	246	66.5	482	19	US-09-509-031-16	Sequence 16, Appli
35	246	66.5	495	19	US-09-509-031-4	Sequence 4, Appli
36	246	66.5	495	23	US-09-820-048A-4	Sequence 4, Appli
37	73	19.7	1447	30	US-10-437-963-198658	Sequence 198658, A
38	71	19.2	166	31	US-60-452-680-16977	Sequence 16977, A
39	70	18.9	212	16	US-09-270-767-32555	Sequence 32555, A
40	70	18.9	212	16	US-09-270-767-47772	Sequence 47772, A
41	70	18.9	212	16	US-09-270-767-185762	Sequence 185762, A
42	69.5	18.8	1032	1	PCT-US03-07858A-1457	Sequence 1457, Ap
43	69.5	18.8	1032	1	PCT-US03-07858A-1457	Sequence 1457, Ap
44	69.5	18.8	1086	1	PCT-US03-07858-2239	Sequence 2239, Ap
45	69.5	18.8	1086	1	PCT-US03-07858A-2239	Sequence 2239, Ap

#### ALIGNMENTS

RESULT 1  
US-09-808-212A-10  
; Sequence 10, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Robertts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-10

Query Match 100.0%; Score 370; DB 23; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5.5e-37;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEKPEVTKANLIYADGKTQTAETKGFATATAYRYADLLAKENKYYTADLE 60  
100.0%|=====|  
370 100.0 72 23 US-09-808-212A-10  
370 100.0 305 7 US-08-325-278-1

```
Db 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTDLE 60
QY 61 DGGYTINIRFAG 72
| | | | | | | | | |
Db 61 DGGYTINIRFAG 72

RESULT 2
US-08-325-278-1
; Sequence 1, Application US/08325278
; GENERAL INFORMATION:
; APPLICANT: Björck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMASTERS, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278-1
Query Match 100.0%; Score 370; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 4.le-36;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTDLE 60
| | | | | | | | | |
Db 225 KKTPEEPKEEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTDLE 284

QY 61 DGGYTINIRFAG 72
| | | | | | | | | |
Db 285 DGGYTINIRFAG 296

RESULT 3
US-08-325-278A-1
; Sequence 1, Application US/08325278A
; GENERAL INFORMATION:
; APPLICANT: Björck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278A-1
Query Match 100.0%; Score 370; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 4.le-36;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTDLE 60
| | | | | | | | | |
Db 225 KKTPEEPKEEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTDLE 284

QY 61 DGGYTINIRFAG 72
| | | | | | | | | |
Db 285 DGGYTINIRFAG 296

RESULT 4
US-08-325-278B-1
; Sequence 1, Application US/08325278B
; GENERAL INFORMATION:
; APPLICANT: Björck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
```

REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 100084.402  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-08-325-278B-1

Query Match 100.0%; Score 370; DB 7; Length 305;  
Best Local Similarity 100.0%; Pred. No. 4.1e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEPKKEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
|||||  
DB 225 KKTPEPKKEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 284  
|||||

QY 61 DGGYTINIRFAG 72  
|||||  
DB 285 DGGYTINIRFAG 296  
|||||

RESULT 5

US-08-325-278-3

; Sequence 3, Application US/08325278  
; GENERAL INFORMATION:  
; APPLICANT: Bjvick, Lars  
; APPLICANT: Sjobring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278  
; FILING DATE: 26-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McWaters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 450023.401  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055

US-08-325-278-3

Query Match 100.0%; Score 370; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 6.7e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEPKKEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
|||||  
DB 225 KKTPEPKKEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 284  
|||||

QY 61 DGGYTINIRFAG 72  
|||||  
DB 285 DGGYTINIRFAG 296  
|||||

RESULT 6

US-08-325-278A-3

; Sequence 3, Application US/08325278A  
; GENERAL INFORMATION:  
; APPLICANT: Bjvick, Lars  
; APPLICANT: Sjobring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278A  
; FILING DATE: 26-Oct-1994  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E. R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 100084.402  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-08-325-278A-3

Query Match 100.0%; Score 370; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 6.7e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEPKKEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
|||||  
DB 225 KKTPEPKKEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 284  
|||||

QY 61 DGGYTINIRFAG 72  
|||||  
DB 285 DGGYTINIRFAG 296  
|||||

RESULT 7

US-08-325-278B-3

; Sequence 3, Application US/08325278B  
; GENERAL INFORMATION:  
; APPLICANT: Bjorck, Lars  
; Sjoerding, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278B  
; FILING DATE: 26-Oct-1994  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E. R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 100084.402  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-08-325-278b-3  
  
Query Match 100.0%; Score 370; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 6.7e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFAEATAEAYRYADLLAKENGKYTDLE 60  
|||||  
Db 225 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFAEATAEAYRYADLLAKENGKYTDLE 284  
|||||  
  
QY 61 DGGYTINIRFAG 72  
Db 285 DGGYTINIRFAG 296  
  
RESULT 8  
US-09-980-469-12  
; Sequence 12, Application US/09980469  
; GENERAL INFORMATION:  
; APPLICANT: ziv, Shani  
; APPLICANT: Shoseyov, Oded  
; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND RECO  
; TITLE OF INVENTION: PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTURED  
; FILE REFERENCE: 01/22924  
; CURRENT APPLICATION NUMBER: US/09/980,469  
; CURRENT FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:

; OTHER INFORMATION: pUC19-cell-Protl-cesNG-HDEL fusion encoded product  
US-09-980-469-12  
  
Query Match 100.0%; Score 370; DB 25; Length 467;  
Best Local Similarity 100.0%; Pred. No. 7.4e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFAEATAEAYRYADLLAKENGKYTDLE 60  
|||||  
Db 265 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFAEATAEAYRYADLLAKENGKYTDLE 324  
|||||  
  
QY 61 DGGYTINIRFAG 72  
Db 325 DGGYTINIRFAG 336  
|||||  
  
RESULT 9  
US-09-791-537-10210  
; Sequence 10210, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10210  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus magnus  
US-09-791-537-10210  
  
Query Match 100.0%; Score 370; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 1.4e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFAEATAEAYRYADLLAKENGKYTDLE 60  
|||||  
Db 318 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFAEATAEAYRYADLLAKENGKYTDLE 377  
|||||  
  
QY 61 DGGYTINIRFAG 72  
Db 378 DGGYTINIRFAG 389  
|||||  
  
RESULT 10  
US-09-791-537-96101  
; Sequence 96101, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 96101  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus magnus  
US-09-791-537-96101  
  
Query Match 100.0%; Score 370; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 1.4e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
Db 318 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 377  
Qy 61 DGGYTINIRFAG 72  
Db 378 DGGYTINIRFAG 389

RESULT 11  
US-09-808-212A-8  
; Sequence 8, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-8

Query Match 91.6%; Score 339; DB 23; Length 72;  
Best Local Similarity 90.3%; Pred. No. 3.4e-33;  
Matches 65; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
Db 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADVA 60  
Qy 61 DGGYTINIRFAG 72  
Db 61 DGGYTINIRFAG 72

RESULT 12  
US-09-808-212A-18  
; Sequence 18, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-18

Query Match 88.9%; Score 329; DB 23; Length 75;  
Best Local Similarity 90.0%; Pred. No. 6e-32;  
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KTEPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLEDG 62  
Db 5 ETPEPEKKEVTIKVNLIFADGKTQTAEFGTFAEATAEAYRYADLLAKVNGEYTDLEDG 64  
Qy 63 GYTINIRFAG 72  
Db 65 GYTINIRFAG 74

RESULT 13  
US-08-446-137A-4  
; Sequence 4, Application US/08446137A  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; FROM PROTEIN L AND THEIR USES  
; NUMBER OF SEQUENCES: 4  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,137A  
; FILING DATE: 22-05-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MITCHARD, LEONARD C.  
; REGISTRATION NUMBER: 29,009  
; REFERENCE/DOCKET NUMBER: 1418-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-446-137A-4

Query Match 88.9%; Score 329; DB 8; Length 291;  
Best Local Similarity 90.0%; Pred. No. 4e-31;  
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KTEPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLEDG 62  
Db 221 ETPEPEKKEVTIKVNLIFADGKTQTAEFGTFAEATAEAYRYADLLAKVNGEYTDLEDG 280  
Qy 63 GYTINIRFAG 72  
Db 281 GYTINIRFAG 290

RESULT 14  
US-09-791-537-88366  
; Sequence 88366, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 88366  
; LENGTH: 992  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus magnus  
US-09-791-537-88366

Query Match 88.9%; Score 329; DB 22; Length 992;  
Best Local Similarity 90.0%; Pred. No. 2.2e-30;  
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KTEPEPKKEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLEDG 62  
Db 468 ETPEPEKKEVTIKVNLIFADGKTQTAEFGTFAEATAEAYRYADLLAKVNGEYTDLEDG 527

QY 63 GYTINIRFAG 72  
| | | | | : | | |  
Db 528 GYTINIKFAG 537

```

RESULT 15
US-08-331-637-2
; Sequence 2, Application US/08331637
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEIN L AND PROCESS FOR ITS PREPARATION BY
; RECOMBINANT DNA TECHNOLOGY
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-331-637-2

```

Query Match 88.9%; Score 329; DB 7; Length 1027;  
Best Local Similarity 90.0%; Pred. No. 2.3e-30;  
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY	3	K	T	P	E	E	P	K	E	E	V	T	I	K	A	N	L	I	Y	A	D	G	K	T	A	E	F	K	G	T	F	A	E	A	T	A	E	A	Y	A	D	L	L	A	K	E	N	G	K	Y	T	A	D	E	G	62	
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
Db	503	E	T	P	E	E	P	K	E	E	V	T	I	K	V	N	L	I	F	A	D	G	K	T	A	E	F	K	G	T	F	E	A	T	A	E	A	Y	A	D	L	L	A	K	V	E	N	G	Y	T	A	D	E	G	562		

Qy 63 GYTINIRFAG 72  
| | | | | : | | |  
Db 563 GYTINIKFAG 572

Search completed: September 3, 2003, 11:40:09  
Job time : 214.041 secs



```
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)E
; CURRENT APPLICATION NUMBER: US/10/613,520
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 10/389,566
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2578
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2239
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-520-2239

Query Match      18.8%; Score 69.5; DB 6; Length 1086;
Best Local Similarity 25.6%; Pred. No. 2.5;
Matches 21; Conservative 15; Mismatches 29; Indels 17; Gaps 1;

QY 7 EPKEEVTIRANLIYADGKTQTAEFKGTAEAT-----AEAYRYADLLA 49
Db 497 EREDEDLGPASFVVSAGQRAASSVGLFAGSSNNIYISTQSOFVNTLSELFQSARLSA 556
QY 50 KENGKYTADLEDGGYTNIRFA 71
Db 557 SSLRYVYGLGLENGYTVTLQFA 578

RESULT 3
US-10-637-544-2
; Sequence 2, Application US/10637544
; GENERAL INFORMATION:
; APPLICANT: Chr. Hansen A/S
; TITLE OF INVENTION: Peptides with anti-hypertensive properties
; FILE REFERENCE: PI0320501
; CURRENT APPLICATION NUMBER: US/10/637,544
; CURRENT FILING DATE: 2003-08-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1849
; TYPE: PRT
; ORGANISM: Lactobacillus helveticus
US-10-637-544-2

Query Match      18.8%; Score 69.5; DB 6; Length 1849;
Best Local Similarity 35.1%; Pred. No. 4.8;
Matches 26; Conservative 9; Mismatches 28; Indels 11; Gaps 4;

QY 1 KEKTPPEPKEEVTIRKANL---IYADGKTQ----TAEFKGTFAEATAEAYRYADLLAKENG 53
Db 1590 KENQPSDEKELAKENLOAKYDAGEKDLDRYTADSKDFNDALKKA---KDVLDKNA 1646
QY 54 KYTADLEDGGYTN 67
Db 1647 K-LADLQDAAKALD 1659

RESULT 4
US-10-613-520-1082
; Sequence 1082, Application US/10613520
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; APPLICANT: Wu, Wei
; APPLICANT: Maolong, Lu
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)E
```

```
; CURRENT APPLICATION NUMBER: US/10/613,520
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 10/389,566
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2578
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1082
; LENGTH: 2062
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-520-1082

Query Match      18.8%; Score 69.5; DB 6; Length 2062;
Best Local Similarity 25.6%; Pred. No. 5.6;
Matches 21; Conservative 15; Mismatches 29; Indels 17; Gaps 1;

QY 7 EPKEEVTIRANLIYADGKTQTAEFKGTAEAT-----AEAYRYADLLA 49
Db 1473 EREDEDLGPASFVVSAGQRAASSVGLFAGSSNNIYISTQSOFVNTLSELFQSARLSA 1532
QY 50 KENGKYTADLEDGGYTNIRFA 71
Db 1533 SSLRYVYGLGLENGYTVTLQFA 1554

RESULT 5
US-10-613-520-1944
; Sequence 1944, Application US/10613520
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; APPLICANT: Wu, Wei
; APPLICANT: Maolong, Lu
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)E
; CURRENT APPLICATION NUMBER: US/10/613,520
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 10/389,566
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2578
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1944
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-520-1944

Query Match      18.5%; Score 68.5; DB 6; Length 858;
Best Local Similarity 35.7%; Pred. No. 2.4;
Matches 20; Conservative 9; Mismatches 22; Indels 5; Gaps 2;

QY 16 ANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTADLEDGGYTNIRFA 71
Db 305 SNNIYI--ATSLAQFINTM---DSELFQSARLSASSRLRYVYGLGLENGYTVTLQFA 355

RESULT 6
US-10-613-520-1237
; Sequence 1237, Application US/10613520
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
```

[illegible]

```
; APPLICANT: Maolong, Lu
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)E
; CURRENT APPLICATION NUMBER: US/10/613,520
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 10/389,566
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2578
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1238
; LENGTH: 1032
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-613-520-1238

Query Match      18.0%; Score 66.5; DB 6; Length 1032;
Best Local Similarity 33.9%; Pred. No. 5.4;
Matches 19; Conservative 10; Mismatches 22; Indels 5; Gaps 2;

QY 16 ANLIYADGKTOTAEFKGTFAEATAEAYRYADLLAKENGYTADLEDGGYTIIRFA 71
Db 476 SNNIYL--ATSOFOVNTL---DSSELFOSARUSASSVRYGLGNGYTVTLQFA 526

RESULT 11
PCT-US03-19070-2
; Sequence 2, Application PC/TUS0319070
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: INHIBITION OF T CELL ACTIVATION BY
; FILE REFERENCE: GNN-042PC
; CURRENT APPLICATION NUMBER: PCT/US03/19070
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 60/389660
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-19070-2

Query Match      17.6%; Score 65; DB 1; Length 334;
Best Local Similarity 30.4%; Pred. No. 2;
Matches 17; Conservative 11; Mismatches 18; Indels 10; Gaps 2;

QY 19 IYADGK----TOTAEFKGTFA-----EATAEAYRYADLLAKENGYTADLEDGGY 64
Db 78 VYADGKEVEDRQSAPYRGRTSILRDGITAGKALRIHNVTASDSGKYLCTFQDGF 133

RESULT 12
US-10-463-260-2
; Sequence 2, Application US/10463260
; GENERAL INFORMATION:
; APPLICANT: Wood, Clive R.
; APPLICANT: Carreno, Beatriz M.
; TITLE OF INVENTION: INHIBITION OF T CELL ACTIVATION BY
; FILE REFERENCE: GNN-042
; CURRENT APPLICATION NUMBER: US/10/463,260
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 60/389660
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-463-260-5

Query Match      17.6%; Score 65; DB 6; Length 481;
Best Local Similarity 30.4%; Pred. No. 3.1;
Matches 17; Conservative 11; Mismatches 18; Indels 10; Gaps 2;

QY 19 IYADGK----TOTAEFKGTFA-----EATAEAYRYADLLAKENGYTADLEDGGY 64
Db 78 VYADGKEVEDRQSAPYRGRTSILRDGITAGKALRIHNVTASDSGKYLCTFQDGF 133

RESULT 13
PCT-US03-19070-5
; Sequence 5, Application PC/TUS0319070
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: INHIBITION OF T CELL ACTIVATION BY
; FILE REFERENCE: GNN-042PC
; CURRENT APPLICATION NUMBER: PCT/US03/19070
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 60/389660
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-19070-5

Query Match      17.6%; Score 65; DB 1; Length 481;
Best Local Similarity 30.4%; Pred. No. 3.1;
Matches 17; Conservative 11; Mismatches 18; Indels 10; Gaps 2;

QY 19 IYADGK----TOTAEFKGTFA-----EATAEAYRYADLLAKENGYTADLEDGGY 64
Db 78 VYADGKEVEDRQSAPYRGRTSILRDGITAGKALRIHNVTASDSGKYLCTFQDGF 133

RESULT 14
US-10-463-260-5
; Sequence 5, Application US/10463260
; GENERAL INFORMATION:
; APPLICANT: Wood, Clive R.
; APPLICANT: Long, Andrew J.
; APPLICANT: Carreno, Beatriz M.
; TITLE OF INVENTION: INHIBITION OF T CELL ACTIVATION BY
; FILE REFERENCE: GNN-042
; CURRENT APPLICATION NUMBER: US/10/463,260
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 60/389660
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-463-260-5

Query Match      17.6%; Score 65; DB 6; Length 481;
Best Local Similarity 30.4%; Pred. No. 3.1;
Matches 17; Conservative 11; Mismatches 18; Indels 10; Gaps 2;

QY 19 IYADGK----TOTAEFKGTFA-----EATAEAYRYADLLAKENGYTADLEDGGY 64
```

```

:||||| |:| |:| : | | | :| :| :| :| :
78 VYADGVEVEDRSAPYRGRTSILRDGTAGKAALRIHNVTASDSGKYLCYFQDGF 133
db

```

```

RESULT 15
US-09-200-650E-7
; Sequence 7, Application US/09200650E
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O. Ni
; APPLICANT: Eidhann, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283U52/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-7

```

	Query Match	16.8%	Score 62	DB 5	Length 1166
	Best Local Similarity	31.9%	Pred. NO. 23		
	Matches	15	Conservative	8	Mismatches 20
					Indels 4
					Gaps 2
QY	24	KTQTAFKGTFFAATAEAYRYADLLAKENGKKTADLE-DGGYITINIR	69		
Db	752	KDSTGKVGIVGTTDASGKYKFTDL---DNGNVYVEETPAGYPTVK	795		

Search completed: September 3, 2003, 11:40:57  
Job time : 11.1096 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:52 ; Search time 9.12329 Seconds  
(without alignments)  
371.130 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_225\_296

Perfect score: 370

Sequence: 1 KKKTPPEPKEVTIRANLIY.....GKYTADLEGGTYTINIRFAG 72

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	71.5	19.3	398	1	RPC4_HUMAN
2	70.5	19.1	398	1	RPC4_MOUSE
3	64.5	17.4	1179	1	DP3A_STRCO
4	64	17.3	2459	1	MAPB_RAT
5	64	17.3	2464	1	MAPB_MOUSE
6	63.5	17.2	367	1	GPR_OCEIH
7	63.5	17.2	550	1	THS_PYRAB
8	62	16.8	953	1	LKAL_PASHA
9	62	16.8	953	1	LKAB_PASHA
10	61.5	16.6	404	1	ASSY_VIBCH
11	61.5	16.6	439	1	GLNA_PYRAB
12	61	16.5	471	1	TNAE_ECO57
13	61	16.5	471	1	TNAE_ECOLI
14	60.5	16.4	238	1	RNPH_VIBCH
15	60.5	16.4	263	1	RCY3_ASTLO
16	60.5	16.4	451	1	Y996_METJA
17	60.5	16.4	556	1	SYR_LISIN
18	60.5	16.4	674	1	COAA_BACTJ
19	60.5	16.4	692	1	MYS_PODCA
20	60.5	16.4	787	1	OSTA_VIBCH
21	60	16.2	451	1	TBB4_PORPU
22	60	16.2	596	1	RGPI_DROME
23	59.5	16.1	245	1	RNPH_STRCO
24	59.5	16.1	257	1	FAEJ_ECOLI
25	59.5	16.1	390	1	IRKB_CAVPO
26	59.5	16.1	425	1	IRK6_MOUSE
27	59.5	16.1	425	1	IRK6_RAT
28	59.5	16.1	658	1	TRAG_AGR5
29	59	15.9	490	1	CPCJ_HUMAN
30	59	15.9	582	1	TNPA_BACFR
31	59	15.9	627	1	ABPX_YEAST
32	58.5	15.8	374	1	FCGL_HUMAN
33	58.5	15.8	423	1	IRK6_HUMAN

#### RESULT 1

ID	RPC4_HUMAN	STANDARD;	PRT;	398 AA.
AC	P05423; Q9BPV7; Q9BP21; Q9BXB3;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	DNA-directed RNA polymerase III 47 kDa polypeptide (EC 2.7.7.6) (RNA polymerase C subunit 4) (RPC4) (RPC53) (BN51 protein).			
GN	BN51T OR BN51 OR POLR3D.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88065472; PubMed=3683386;			
RA	Ittmann M., Greco A., Basilico C.;			
RT	"Isolation of the human gene that complements a temperature-sensitive cell cycle mutation in BHK cells."			
RL	Mol. Cell. Biol. 7:3386-3393(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22278866; PubMed=12391170;			
RA	Hu P., Wu S., Sun Y., Yuan C.-C., Kobayashi R., Myers M.P., Hernandez N.;			
RT	"Characterization of human RNA polymerase III identifies orthologues for Saccharomyces cerevisiae RNA polymerase III subunits."			
RL	Mol. Cell. Biol. 22:8044-8055(2002).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung, and Lymph;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L., Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[4]			
RP	SEQUENCE OF 1-50 FROM N.A.			
RX	MEDLINE=21282972; PubMed=11279001;			

P49658 mesocricetu  
P36049 saccharomyc  
P33536 osmunda cin  
P22346 lactococcus  
Q94526 drosophila  
P05519 bacillus th  
Q8Y7N6 listeria mo  
Q58367 methanococ  
Q8UHN3 agrobacteri  
P48070 euglena gen  
Q94740 schizosacch  
P00878 euglena gra

#### ALIGNMENTS



polymerase III complex (By similarity).  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -|- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. DNAAE  
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC EMBL; AF108191; AAD16978.1; -  
 CC EMBL; AL096884; CAB51456.1; -  
 CC PIR; T35093; T35093.  
 CC InterPro: IPR004013; PHP.C.  
 CC InterPro: IPR003141; PHP\_N.  
 CC InterPro: IPR004805; PHP\_alpha.  
 CC InterPro: IPR004365; trna\_anti.  
 CC Pfam; PF02811; PHP\_C; 1.  
 CC Pfam; PF02231; PHP\_N; 1.  
 CC SMART; SM00481; POLIIIAC; 1.  
 CC TIGRFAMS; TIGR00594; polc; 1.  
 CC Transferase: DNA-directed DNA polymerase; DNA replication;  
 CC Complete proteome.  
 CC KW TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION;  
 CC FT VARIANT 802 802 E -> K (IN MUTANT TS-38).  
 CC SQ SEQUENCE 1179 AA; 130795 MW; 7E4B58675B634CD3 CRC64;

Query Match 17.4%; Score 64.5; DB 1; Length 1179;  
 Best Local Similarity 31.5%; Pred. No. 40;  
 Matches 23; Conservative 9; Mismatches 34; Indels 7; Gaps 2;  
 FT VARIATION 802 802 E -> K (IN MUTANT TS-38).  
 SQ SEQUENCE 1179 AA; 130795 MW; 7E4B58675B634CD3 CRC64;

QY 2 EKTPEEPKEEVT-----TKANLIYADGKTQTAEFGTAEATAEAYRYADLLAKENGYT 56  
 Db 131 QKWPRMDKETISQWSEGVASTGCPSEGVOTRLRLGHFEDEALKAAADYQDIFGKD--RYF 188

QY 57 ADLEGGYTNIR 69  
 Db 189 LEIMDHGIEHR 201

RESULT 4  
 MAPB\_RAT  
 ID MAPB\_RAT STANDARD; PRT; 2459 AA.  
 AC P15205; Q62958; Q9ER21; Q9QW92;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Microtubule-associated protein 1B (MAP1B) (Neuraxin) [Contains: MAP1  
 DE light chain LC1].  
 GN MAP1B.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=101116;  
 RN [1]  
 RP SEQUENCE OF 1-142 FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;  
 RX MEDLINE=96257242; PubMed=8666295;  
 RA Liu D., Fischer I.;  
 RT "Isolation and sequencing of the 5' end of the rat microtubule-  
 RL associated protein (MAP1B)-encoding cDNA";  
 RL Gene 172:307-308(1996).  
 RN [2]  
 RP SEQUENCE OF 96-2459 FROM N.A., DOMAIN, AND INDUCTION.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain, and Glial tumor;  
 RX MEDLINE=92347374; PubMed=1639092;  
 RA Zauner W., Kratz J., Staunton J., Feick P., Wiche G.;  
 RT "Identification of two distinct microtubule binding domains on  
 RL recombinant rat MAP 1B";  
 RL Eur. J. Cell Biol. 57:66-74(1992).

[3]  
 RP SEQUENCE OF 1541-2459 FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=spinal cord;  
 RX MEDLINE=90059871; PubMed=2551510;  
 RA Rientitz A., Grenningloh G., Hermans-Borgmeyer I., Kirsch J.,  
 RA Littauer U.Z., Prior P., Gundelfinger E.D., Schmitt B., Betz H.;  
 RT "Neuraxin, a novel putative structural protein of the rat central  
 RT nervous system that is immunologically related to microtubule-  
 RT associated protein 5.";  
 RL EMBO J. 8:2879-2888(1989).  
 [4]  
 RP DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.  
 RX MEDLINE=97405699; PubMed=9260743;  
 RA Ma D., Nothias F., Boyne L.J., Fischer I.;  
 RT "Differential regulation of microtubule-associated protein 1B (MAP1B)  
 RT in rat CNS and PNS during development.";  
 RL J. Neurosci. Res. 49:319-332(1997).  
 CC -1- FUNCTION: The function of brain MAPS is essentially unknown.  
 CC Phosphorylated MAP1B may play a role in the cytoskeletal changes  
 CC that accompany neurite extension. Possibly MAP1B binds to at least  
 CC two tubulin subunits in the polymer, and this bridging of subunits  
 CC might be involved in nucleating microtubule polymerization and in  
 CC stabilizing microtubules.  
 CC -1- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate  
 CC with MAP1A and MAP1B proteins.  
 CC -1- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem,  
 CC cerebellum and cerebrum). Not expressed in liver, spleen, kidney,  
 CC heart or muscle.  
 CC -1- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic  
 CC nerve levels are high early in development but decrease during  
 CC postnatal development and are low in adults. In dorsal root  
 CC ganglia levels remain high throughout development.  
 CC -1- INDUCTION: By nerve growth factor.  
 CC -1- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC responsible for the binding of MAP1B to microtubules.  
 CC -1- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated  
 CC from MAP1B by proteolytic processing. It is free to associate with  
 CC both MAP1A and MAP1B. It interacts with the amino-terminal region  
 CC of MAP1B (by similarity).  
 CC -1- PTM: Phosphorylated.  
 CC -1- SIMILARITY: TO MAP1A.  
 CC -1- CAUTION: A C-terminal fragment of this protein (residues 1597 to  
 CC 2459) was originally described as neuraxin in Ref.3.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; U52950; AAB17068.1;  
 DR EMBL; X60370; CAC16162.1;  
 DR EMBL; X16623; CAA34620.1; ALT\_SEQ.  
 DR PIR; A56577; A56577;  
 DR InterPro; IPR000102; MAP1B\_neuraxin.  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 8.  
 KW Microtubules; Repeat; Phosphorylation.  
 FT CHAIN ? 2459 MAP1 LIGHT CHAIN LC1.  
 FT REPEAT 1869 1885 MAP1B 1.  
 FT REPEAT 1886 1902 MAP1B 2.  
 FT REPEAT 1903 1919 MAP1B 3.  
 FT REPEAT 1920 1936 MAP1B 4.  
 FT REPEAT 1937 1953 MAP1B 5.  
 FT REPEAT 1954 1970 MAP1B 6.  
 FT REPEAT 1988 2004 MAP1B 7.  
 FT REPEAT 2005 2021 MAP1B 8.  
 FT REPEAT 2022 2038 MAP1B 9.  
 FT REPEAT 2039 2055 MAP1B 10.  
 FT DOMAIN 559 1035 GLU-RICH.

FT DOMAIN 588 786 LYS-RICH (HIGHLY BASIC, CONTAINS MANY  
 FT KKEE AND KKEI/V REPEATS).  
 FT DOMAIN 2224 2312 LYS-RICH.  
 FT CONFLICT 127 127 M -> V (IN REF. 1).  
 FT CONFLICT 140 140 T -> S (IN REF. 1).  
 FT CONFLICT 2112 2112 R -> K (IN REF. 3).  
 FT CONFLICT 2169 2169 L -> I (IN REF. 3).  
 SQ SEQUENCE 2459 AA; 269497 MW; 2E3F6872DED8BA2 CRC64;  
 Query Match 17.3%; Score 64; DB 1; Length 2459;  
 Best Local Similarity 30.4%; Pred. No. 1e+02;  
 Matches 24; Conservative 11; Mismatches 26; Indels 18; Gaps 4;  
 QY 2 EKTPEPKK-----EVTIKANLIYADG-----KTQAEFGKGTFAETATAYADL 47  
 DB 1918 EKTTPKPDGGYSCETETRTPEGGYSYEISEKTRTPPEVSGTYETKTERSRLLDD 1977  
 QY 48 LAKENKYYTADLEDGGYTI 66  
 DB 1978 IS--NG--YDDEGGHTL 1992  
 RESULT 5  
 MAP1B MOUSE  
 ID MAP1B\_MOUSE STANDARD; PRT; 2464 AA.  
 AC F14873;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Microtubule-associated protein 1B (MAP 1B) (MAP1.2) (MAP1(X))  
 DE [Contains: MAP1 light chain LC1].  
 GN MAP1B OR MTAP1B OR MTAP5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND DOMAIN.  
 RC STRAIN=Swiss Webster; TISSUE=Brain;  
 RX MEDLINE=90094539; PubMed=2480963;  
 RA Noble M., Lewis S.A., Cowan N.J.;  
 RT "The microtubule binding domain of microtubule-associated protein  
 RT and tau.";  
 RT MAP1B contains a repeated sequence motif unrelated to that of MAP2  
 RL J. Cell Biol. 109:3367-3376(1989).  
 CC -1- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.  
 CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES  
 CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST  
 CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS  
 CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN  
 CC STABILIZING MICROTUBULES.  
 CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE  
 CC WITH MAP1A AND MAP1B PROTEINS.  
 CC -1- DOMAIN: Has a highly basic region with many copies of the sequence  
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is  
 CC responsible for the binding of MAP1B to microtubules.  
 CC -1- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED  
 CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH  
 CC BOTH MAP1A AND MAP1B. IT INTERACTS WITH THE AMINO-TERMINAL REGION  
 CC OF MAP1B.  
 CC -1- SIMILARITY: TO MAP1A.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; X51396; CAA35761.1;  
 DR PIR; S07549; QRMSP1.  
 DR MGD; MGI:1306778; Mtap1b.

PFam: PF03418; Peptidase\_U3; 1.  
TIGRfams: TIGR01441; GPR; 1.  
KW Hydrolyase; Protease; Zymogen; Complete proteome.  
FT PROPEP 1 13 BY SIMILARITY.  
FT CHAIN 14 367 GERMINATION PROTEASE.  
SQ SEQUENCE 367 AA; 40467 MW; 3F4985F73F2CE5C7 CRC64;

Query Match 17.2%; Score 63.5; DB 1; Length 367;  
Best Local Similarity 28.0%; Pred. No. 14;  
Matches 21; Conservative 14; Mismatches 27; Indels 13; Gaps 4;

QY 1 KEKTPPEKPEVTKANLIYADGK---TQTAEKKG-TFAEATAEAYR--YAD-----L 47  
Db 3 EQQIPQVRYDLAIEAKDMYTESKPEETNDKEIKGVTFKRSVKDIKVSVDIDEGEKL 62  
QY 48 LAKENGKYTADLEDG 62  
Db 63 LGKKPGSVYIYADG 77

RESULT 7  
THS\_PYRAB  
ID THS\_PYRAB STANDARD; PRT; 550 AA.  
AC Q2V2Q7;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Thermosome subunit (Chaperonin subunit).  
GN THS OR THSA OR PYRAB00180 OR PAB2341.  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=29292;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GE5 / Orsay;  
RX PubMed=12622808;  
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,  
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,  
RA Van der Oost J., Weissbach J., Zivanovic Y., Forterre P.;  
RT "An integrated analysis of the genome of the hyperthermophilic  
archaeon Pyrococcus abyssi".  
RL Mol. Microbiol. 47:1495-1512(2003).  
CC -1- FUNCTION: MOLECULAR CHAPERONE; BINDS UNFOLDED POLYPEPTIDES IN  
CC VITRO, AND HAS A WEAK ATPASE ACTIVITY (BY SIMILARITY).  
CC -1- SUBUNIT: FORMS AN OLIGOMERIC COMPLEX OF EIGHT-MEMBERED RINGS  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: Belongs to the TCP-1 chaperonin family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: AJ248283; CAB48941.1; -  
CC PIR: F75186; F75186.  
CC HSSP: P48424; 1A5S.  
CC InterPro: IPR002194; Chaperonin\_TCP-1.  
CC InterPro: IPR001844; Chaperonin\_TCP60.  
CC InterPro: IPR002423; Cpn60/TCP-1.  
CC Pfam: PF00118; cpn60\_TCP1\_1.  
CC PRINTS: PR00298; CHAPERONIN60.  
CC PRINTS: PR00304; TCOMPLEXTCP1.  
CC PROSITE: PS00750; TCP1\_1; 1.  
CC PROSITE: PS00751; TCP1\_2; 1.  
CC PROSITE: PS00995; TCP1\_3; 1.  
CC Chaperone; ATP-binding; Complete proteome.  
KW SEQUENCE 550 AA; 59717 MW; 523196D4B52BAC53 CRC64;  
SQ

Query Match 17.2%; Score 63.5; DB 1; Length 550;

Best Local Similarity 31.5%; Pred. No. 22;  
Matches 23; Conservative 9; Mismatches 24; Indels 17; Gaps 3;

QY 1 KEKTPPEKKEVITKANLIYADGKTQTAEFGKFAEATAEAYRYADLLAKENGKYTDLE 60  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 144 KEVKPDD--EVLKKAAMTAITGK-----AAEEERYLAKLAVEAVKLVAEEK 189

QY 61 DGGYTI---NIRE 70  
|| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 190 DGRKFKVDIDNIKF 202

RESULT 8  
LKAL\_PASHA STANDARD; PRT; 953 AA.  
AC P16535;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leukotoxin from serotype A1.  
GN LKTA.  
OS Pasteurella haemolytica.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Mannheimia.  
OX NCBI\_TaxID=75985;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serotype A1.  
RX MEDLINE=87306837; PubMed=3040588;  
RA Lo R.Y.C.; Strathdee C.A.; Shewen P.E.;  
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella  
haemolytica A1."  
RL Infect. Immun. 55:1987-1996(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serotype A1 / PH1101;  
RX MEDLINE=89210283; PubMed=2707120;  
RA Highlander S.K.; Chidambaram M.; Engler M.J.; Weinstock G.M.;  
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene  
cluster."  
RL DNA 8:15-28(1989).  
RN [3]  
RP SEQUENCE OF 884-953 FROM N.A.  
RC STRAIN=Serotype A1 / PH1101;  
RX MEDLINE=90236888; PubMed=2185213;  
RA Highlander S.K.; Engler M.J.; Weinstock G.M.;  
RT "Secretion and expression of the Pasteurella haemolytica Leukotoxin."  
RL J. Bacteriol. 172:2343-2350(1990).  
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood  
cell membranes and cause cell rupture by mechanisms not clearly  
defined.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: The Gly-rich region is probably involved in binding  
calcium, which is required for target cell-binding or cytolytic  
activity.  
CC -!- DOMAIN: The three transmembrane domains are believed to be  
involved in pore formation by the cytotoxin (By similarity).  
CC -!- PTM: Palmitoylated by lktC. The toxin only becomes active when  
modified (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; M20730; AAA25529.1; -  
CC EMBL; M24197; AAA25543.1; -  
CC PIR; B30169; B30169.  
CC InterPro; IPR001343; Hemlysn\_Ca\_bind.

DR InterPro; IPR003995; RTxA  
DR Pfam; PF00353; hemolysinCabin; 5.  
DR Pfam; PF02382; RTX; 1.  
DR PRINTS; PRO0313; CABNDNGRPT.  
DR PRINTS; PRO1488; RTXTOXINA.  
DR PROSITE; PS00330; HEMOLYSIN\_CALCIIUM; 4.  
KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;  
KW Lipoprotein; Palmitate.  
FT TRANSMEM 229 249 POTENTIAL.  
FT TRANSMEM 297 318 POTENTIAL.  
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH (BY SIMILARITY).  
FT REPEAT 734 739 1.  
FT REPEAT 743 748 2.  
FT REPEAT 752 757 3.  
FT REPEAT 761 766 4.  
FT REPEAT 770 775 5.  
FT REPEAT 779 784 6.  
FT CONFLICT 409 414 FEHVN -> LSTLQI (IN REF. 2).  
FT CONFLICT 742 743 D -> Y (IN REF. 2).  
SQ SEQUENCE 953 AA; 101996 MW; 7F93D113A118C05F CRC64;  
Query Match 16.8%; Score 62; DB 1; Length 953;  
Best Local Similarity 33.3%; Pred. No. 59;  
Matches 20; Conservative 7; Mismatches 23; Indels 10; Gaps 1;

QY 9 KEETIKANLIYADGKTQTAEFGKFAEATAE-----AYRYADLLAKENGKYTD 58  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 832 KEKVTIQNFREDAFAKEVPNYKATKDEKIEIIGQGERITSQVDDLLIAGKNGKITD 891

RESULT 9  
LKAL\_PASHA STANDARD; PRT; 953 AA.  
AC P55118;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leukotoxin from serotype A1.  
GN LKTA.  
OS Pasteurella haemolytica.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Mannheimia.  
OX NCBI\_TaxID=75985;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serotype A1;  
RX MEDLINE=94041617; PubMed=8225575;  
RA Burrows L.L.; Olah-Winfield E.; Lo R.Y.C.;  
RT "Molecular analysis of the leukotoxin determinants from Pasteurella  
haemolytica serotypes 1 to 16."  
RL Infect. Immun. 61:5001-5007(1993).  
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood  
cell membranes and cause cell rupture by mechanisms not clearly  
DEFINED.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: The Gly-rich region is probably involved in binding  
calcium, which is required for target cell-binding or cytolytic  
activity.  
CC -!- DOMAIN: The three transmembrane domains are believed to be  
involved in pore formation by the cytotoxin (BY SIMILARITY).  
CC -!- PTM: Palmitoylated by lktC. The toxin only becomes active when  
modified (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; M20730; AAA25529.1; -  
CC EMBL; M24197; AAA25543.1; -  
CC PIR; B30169; B30169.  
CC InterPro; IPR001343; Hemlysn\_Ca\_bind.

```

EMBL; AE004330; AAF95783.1; -
DR DR PIR; A82052; A82052.
DR DR TIGR; VC2642; -.
DR DR HAMAP; MF_00005; -.
DR DR InterPro; IPR001516; Arginosuc_synth.
DR DR Pfam; PF00764; Arginosuc_synth; 1.
DR DR ProDom; PD003544; Arginosuc_synth; 1.
DR DR TIGRFAMS; TIGR00032; argG; 1.
DR DR PROSITE; PS00564; ARGININOSUCCIN_SYN_1; 1.
DR DR PROSITE; PS00565; ARGININOSUCCIN_SYN_2; 1.
DR DR Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
KW KW Arginine 404 AA; 44465 MW; A65C73B442B5AF82 CRC64;
SQ

Query Match 16.6% Score 61.5; DB 1; Length 404;
Best Local Similarity 35.5%; Pred. No. 26;
Matches 22; Conservative 9; Mismatches 22; Indels 9; Gaps 2;

QY 2 EKTPPEPKKEVTKANLIYADGKQTQTAEKCTFAETAETAEVRYADLLAKE-NGKYTDLE 60
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
298 DKTSFEFRELGIKASHLVYDGRWFTPLRQAVFAA-----ADELAKDVNGEYVIKLY 349

QY 61 DG 62
Db 350 KG 351

RESULT 11
GLNA_PYRAB STANDARD; PRT; 439 AA.
ID GLNA_PYRAB
AC Q90Y99;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).
GN GLNA OR PYRAB16090 OR PAB1292.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
CN NCBI_TaxID=29292;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GE5 / Orsay;
RX PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Weissenbach J., Zivanovic V., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RL Mol. Microbiol. 47:1495-1512(2003).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ248288; CAB50513.1; -.
DR PIR; C75009; C75009.
DR HSP; P06201; ILGR.
DR InterPro; IPR001691; GLN_synth.
DR InterPro; IPR004809; GlnA.
DR InterPro; IPR001637; GlnA_adenyltn.
DR Pfam; PF00120; gln-synt; 1.
DR Pfam; PF03951; gln-synt_N; 1.
DR ProDom; PD001057; Gln_synt_C; 1.
DR TIGRFAMS; TIGR00653; GlnA; 1.

```



DR HAMAP: MF\_00544; -; 1.  
DR InterPro: IPR001597; Beta\_elim\_lyase.  
DR Pfam: PF01212; Beta\_elim\_lyase; 1.  
DR PROSITE: PD005927; Beta\_elim\_lyase; 1.  
DR PROSITE: PS00853; BETA\_ELIM\_LYASE; 1.  
KW Tryptophan catabolism; Lyase; Pyridoxal phosphate; Complete proteome  
FT BINDING 270 270  
FT BLINDING 270 270  
FT SEQUENCE 471 AA; 52789 MW; 49DD2D41BD9D0034 CRC64;  
  
Query Match 16.5%; Score 61; DB 1; Length 471;  
Best Local Similarity 41.5%; Pred. No. 35;  
Matches 17; Conservative 7; Mismatches 15; Indels 2; Gaps  
  
QY 24 KTQTAEFKG-TFAETAAYRYADLLAKENGKKTADLEDGG 63  
| : | | : | : | : | | | | | : | : | : | |  
DB 239 KOREABYKDWITQIITRETYKYADMLAM-SAKKDMVPMGG 278  
  
RESULT 13  
TNAA\_ECOLI STANDARD; PRT; 471 AA.  
ID TNAA\_ECOLI STANDARD; PRT; 471 AA.  
AC P00913; P78123;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).  
GN TNAA OR IND OR B3708 OR SF3754.  
OS Escherichia coli, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12;  
RX MEDLINE=82007678; PubMed=6268608;  
RA Deeley M.C., Yanofsky C.;  
RT "Nucleotide sequence of the structural gene for tryptophanase of  
RT Escherichia coli K-12";  
RT J. Bacteriol. 147:787-796(1981).  
RN [2]  
RP SEQUENCE FROM N.A., AND IMPORTANCE OF CYS-298.  
RC SPECIES=E.coli; STRAIN=B/1t7-A;  
RX MEDLINE=89323226; PubMed=2502187;  
RA Tokushige M., Tsujimoto N., Oda T., Honda T., Yumoto N., Ito S.,  
RA Yamamoto M., Kim E.H., Hiragi Y.;  
RT "Role of cysteine residues in tryptophanase for monovalent cation-  
RT induced activation";  
RT Biochimie 71:711-720(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
RX MEDLINE=93315143; PubMed=7686882;  
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;  
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli  
RT genome: organizational symmetry around the origin of replication.";  
RT Genomics 16:551-561(1993).  
RN [4]  
RP SEQUENCE OF TRYPTIC PEPTIDES.  
RC SPECIES=E.coli; STRAIN=K12;  
RX MEDLINE=72134434; PubMed=4551944;  
RA Kagamiyama H., Matsubara H., Snell E.E.;  
RT "The chemical structure of tryptophanase from Escherichia coli. 3.  
RT Isolation and amino acid sequence of the tryptic peptides.";  
RN J. Biol. Chem. 247:1576-1586(1972).  
RN [5]  
RP SEQUENCE OF 1-21 FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12;  
RX MEDLINE=86033634; PubMed=3902796;  
RA Stewart V., Yanofsky C.;  
RT "Evidence for transcription antitermination control of tryptophanase  
RT operon expression in Escherichia coli K-12";  
RN J. Bacteriol. 164:731-740(1985).  
RN [6]



```

RN RP SEQUENCE OF 463-471 FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=91216998; PubMed=2022620;
RA Sarsero J.P., Wooley P.J., Gollnick P.D., Yanofsky C., Pittard A.J.;
RT "A new family of integral membrane proteins involved in transport of
RL aromatic amino acids in Escherichia coli.";
RT J. Bacteriol. 173:3231-3234(1991).
RN RP SEQUENCE OF 1-12.
RC SPECIES=E.coli; STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RL in the genome of Escherichia coli K-12.";
RN RP Electrophoresis 18:1259-1313(1997).
RN RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RX Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RN through comparison with genomes of Escherichia coli K12 and O157.";
RN Nucleic Acids Res. 30:4432-4441(2002).
RN RP MUTAGENESIS OF CYS-294 AND CYS-298.
RC SPECIES=E.coli;
RX MEDLINE=89278130; PubMed=2659590;
RA Phillips R.S., Gollnick P.D.;
RT "Evidence that cysteine 298 is in the active site of tryptophan
RN indole-lyase.";
RL J. Biol. Chem. 264:10627-10632(1989).
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Tryptophan catabolism.
CC -1- SUBUNIT: Homotetramer.
CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K00032; AAA24676.1; -.
DR EMBL; X15974; CA34096.1; -.
DR EMBL; M11990; AAA24679.1; -.
DR EMBL; M59914; -. NOT ANNOTATED_CDS.
DR EMBL; L10328; AAA62059.1; ALT_INIT.
DR EMBL; AE000448; AAC76731.1; ALT_INIT.
DR EMBL; AE015382; AAM45197.1; ALT_INIT.
DR PIR; B65173; WZEC.
DR HSP; P28796; 1AX4.
DR ECO2DBASE; G046.5; 6TH EDITION.
DR Ecogene; Egl1005; tnaA.
DR HAMAP; MF_00544; -. 1.
DR InterPro; IPR001597; Beta_elim_lyase.
DR Pfam; PF01212; Beta_elim_lyase; 1.
DR ProDom; PD005927; Beta_elim_lyase; 1.
DR PROSITE; PS00853; BETA_ELIM_LYASE; 1.
RW Tryptophan catabolism; Lyase; Pyridoxal phosphate; Complete proteome.
FT BINDING 270 270 PYRIDOXAL PHOSPHATE.
FT MUTAGEN 294 294 C->S: IDENTICAL TO WILD-TYPE.
FT MUTAGEN 298 298 C->S: ALTERED ACTIVITY.
FT CONFLICT 137 140 DTTQ -> TTQG (IN REF. 1).
FT CONFLICT 379 380 QA -> TG (IN REF. 1 AND 2).

SQ SEQUENCE 471 AA; 52773 MW; 5AFC1F41BD9D0034 CRC64;
Query Match 16.5%; Score 61; DB 1; Length 471;
Best Local Similarity 41.5%; Pred. No. 35;
Matches 17; Conservative 7; Mismatches 15; Indels 2; Gaps 2;

QY 24 KTQTAEFGK-TFAETAAYRYADLLAKENGKYTADLEDGG 63
DB 239 KOREAEYKDWITQITRETYKYADMLAM-SAKKDAVMPMGG 278
| : ||: | : | : ||: ||: | : | : ||

RESULT 14
RNPH_VIBCH STANDARD; PRT; 238 AA.
AC Q9KVD6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease PH (EC 2.7.7.56) (RNase PH) (tRNA
DE nucleotidyltransferase).
GN RPH OR VC0210.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: Phosphorylatic exoribonuclease that removes nucleotide
CC residues following the -CCA terminus of tRNA and adds nucleotides
CC to the ends of RNA molecules by using nucleoside diphosphates as
CC substrates (By similarity).
CC -1- CATALYTIC ACTIVITY: {tRNA}(N+1) + phosphate -> {tRNA}(N) + a
CC nucleoside diphosphate.
CC -1- SIMILARITY: BELONGS TO THE RNASE PH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004111; AAF93386.1; ALT_INIT.
DR TIGR; VC0210; -. 1.
DR HAMAP; MF_00564; -. 1.
DR InterPro; IPR001247; 3_ExoRNase.
DR InterPro; IPR002381; RNase_PH.
DR Pfam; PF01138; RNase_PH; 1.
DR Pfam; PF03725; RNase_PH_C; 1.
DR PROSITE; PS01277; RIBONUCLEASE_PH; 1.
RW Transferase; Nucleotidyltransferase; tRNA processing;
KW Complete proteome.
SQ SEQUENCE 238 AA; 25631 MW; C07828FBBC900E0D CRC64;
Query Match 16.4%; Score 60.5; DB 1; Length 238;
Best Local Similarity 38.5%; Pred. No. 19;
Matches 15; Conservative 8; Mismatches 13; Indels 3; Gaps 1;

QY 12 VTIKANLIYADGKTQTAEFGKTFATAEAYRYADLLAK 50
| : : || || || || | : | : | : || || || || ||

```

```
Db      112 ITVDCDVIQADGGTRTASITGA---SVALADAFAPHLIAK 147
RESULT 15.
YCY3_LASTLO
ID YCY3_ASTLO STANDARD; PRT; 263 AA.
AC P58147;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 32.5 kDa protein in rpl14-rpl12 intergenic region
DE (ORF263).
OS Astasia longa (Euglenophyceae alga).
OG Chloroplast.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCAP 1204-17a;
RX MEDLINE=21080522; PubMed=11212895;
RA Gockel G., Hachtel W.;
RT "Complete gene map of the plastid genome of the nonphotosynthetic
RT euglenoid flagellate Astasia longa.";
RL Protist 151:347-351(2000).
CC 1- SIMILARITY: BELONGS TO THE A.LONGA ORF167/ORF288 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AJ294725; CAC24603.1;
DR InterPro; IPR006851; DUF613.
DR Pfam; PF04764; DUF613; 2.
KW Hypothetical protein; Chloroplast.
SQ SEQUENCE 263 AA; 32473 MW; 3508B02DB49CB4E6 CRC64;
Query Match 16.4%; Score 60.5; DB 1; Length 263;
Best Local Similarity 31.9%; Pred. No. 21;
Matches 23; Conservative 9; Mismatches 27; Indels 13; Gaps 3;
QY 1 KEKTPPEPKPEVTIKANLIYADGKTQTAEFGKTFEAATAEAYRYADLLAKENGKYTDLE 60
Db 173 KDKYLEEPEPELEYIVLDLF-----LLGPLEKEYSSENYKIYILINERKRY-KDMY 221
QY 61 DGG--YTINIRF 70
Db 222 DMGEIYGINIRF 233
```

Search completed: September 3, 2003, 11:20:07  
Job time : 10.1233 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:17 ; Search time 51.0137 Seconds  
(without alignments)  
236.470 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEETPETPDSSEVETIKR.....GEYTDVADKGYTLNIKIFAG 76

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	100.0	76	21	AAV82537
2	389	100.0	305	14	AAV82537
3	389	100.0	434	14	AAV82537
4	384	98.7	467	22	AAV82537
5	326	83.8	182	18	AAV82537
6	326	83.8	182	20	AAV82537
7	323	83.0	178	18	AAV82537
8	323	83.0	178	20	AAV82537
9	323	83.0	198	18	AAV82537

10	323	83.0	198	20	AAV06910
11	323	83.0	342	18	AAW32482
12	323	83.0	342	18	AAW06909
13	323	83.0	482	20	AAW06915
14	323	83.0	495	18	AAW32480
15	323	83.0	495	20	AAV06908
16	314	80.7	72	21	AAV82538
17	314	80.7	367	21	AAV82538
18	299	76.9	72	21	AAV82539
19	298.5	76.7	1027	14	AAV82539
20	298.5	76.7	1027	14	AAV82539
21	298	76.6	71	21	AAV82541
22	298	76.6	291	14	AAV82541
23	284.5	73.1	75	21	AAV82544
24	270.5	69.5	74	21	AAV82543
25	269	69.2	72	21	AAV82540
26	262.5	67.5	82	21	AAV82536
27	262	67.4	71	21	AAV82542
28	255.5	65.7	82	21	AAV82545
29	255.5	65.7	82	21	AAV82548
30	253.5	65.2	82	21	AAV82546
31	249.5	64.1	82	21	AAV82547
32	71.5	18.4	841	22	AAU34283
33	71.5	18.4	841	22	AAU37158
34	71.5	18.4	930	20	AAV08641
35	71.5	18.4	947	24	ABJ18940
36	69	17.7	123	19	AAW65763
37	69	17.7	1377	22	ABW65439
38	67	17.2	571	12	AAV10376
39	67	17.2	574	23	ABE29074
40	67	17.2	588	22	AAV849638
41	66	17.0	447	22	AAV82903
42	66	17.0	450	22	AAV83066
43	65	16.7	419	23	ABP38596
44	64.5	16.6	280	21	AAV83953
45	64.5	16.6	282	21	AAV09975

#### ALIGNMENTS

RESULT 1  
AAV82537  
ID AAV82537 standard; Protein; 76 AA.  
XX AC AAV82537;  
XX DT 20-JUL-2000 (first entry)  
XX DE Peptostreptococcus strain 312 protein L domain B1 protein sequence.  
XX DE Immunoglobulin light chain binding protein; PPL; protein L;  
XX KW Peptostreptococcus; human immunoglobulin kappa chain;  
XX KW Immunofluorescence chromatography.  
XX OS Peptostreptococcus sp.  
XX PN WO200015803-A1.  
XX PD 23-MAR-2000.  
XX PF 14-SEP-1999; 99WO-GB03048.  
XX PR 14-SEP-1998; 98GB-0019998.  
XX PR 26-APR-1999; 99GB-0009578.  
XX PA (ACTI-) ACTINOVIA LTD.  
XX PI Gore MG, Beckingham JA, Roberts SE;  
XX WPI: 2000-271441/23.  
XX DR N-PSDB; AAA08426.

Amino acid sequenc  
Growth factor TLHL  
TLHL amino acid se  
Amino acid sequenc  
Growth factor CATF  
CATF-TEV aminoaci  
Peptostreptococcus  
Expression vector  
Peptostreptococcus  
Protein L. Peptoc  
Protein L. Peptoc  
Peptostreptococcus  
Immunoglobulin bin  
Peptostreptococcus  
Peptostreptococcus  
Peptostreptococcus  
Ig light chain bin  
Peptostreptococcus  
PPL mutant protein  
PPL mutant protein  
PPL mutant protein  
Staphylococcus aur  
Staphylococcus aur  
S. aureus SdrC pro  
Pathogen specific  
C. elegans DPK-30  
Drosophila melanog  
Streptolysin O der  
Streptococcus poly  
Escherichia coli H  
S. epidermidis ope  
S. epidermidis ope  
Staphylococcus epi  
Gene 15 human secr  
Arabidopsis thalia

PT New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
XX more at pH8 with respect to human immunoglobulin kappa-chain -  
XX  
PS Disclosure; Page 40-41; 56pp; English.  
XX  
CC The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (PI) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
XX Sequence 76 AA;  
Query Match 100.0%; Score 389; DB 21; Length 76;  
Best Local Similarity 100.0%; Pred. No. 1.7e-37;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPETPTDSEEEVTIKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKDNGEY 60  
Db 1 KEETPETPTDSEEEVTIKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKDNGEY 60  
QY 61 VDVAADKGYTLNKFAG 76  
Db 61 VDVAADKGYTLNKFAG 76  
RESULT 2  
AAR42993  
ID AAR42993 standard; Protein; 305 AA.  
XX AC AAR42993;  
XX  
XX 25-MAR-2003 (updated)  
DT 16-MAY-1994 (first entry)  
XX  
DE Immunoglobulin light chain binding protein (Protein L).  
XX  
KW Immunoglobulin; light chain; binding; identification; purification;  
KW separation.  
XX  
OS E. coli LE392/pHDL, DSM 7054.  
XX  
FH Key Location/Qualifiers  
FT Domain 5..305  
FT /label= B1 immunoglobulin light chain binding  
FT domain.  
FT Domain 81..305  
FT /label= B2 immunoglobulin light chain binding  
FT domain.  
FT Domain 153..305  
FT /label= B3 immunoglobulin light chain binding  
FT domain.  
FT Domain 225..305  
FT /label= B4 immunoglobulin light chain binding  
FT domain.  
FT Domain 297..305  
FT /label= B5 immunoglobulin light chain binding  
FT domain.  
XX W09322342-A1.  
XX  
XX 11-NOV-1993.  
XX  
XX 28-APR-1993; 93WO-SE00375.  
XX  
XX 28-APR-1992; 92SE-0001331.  
XX  
XX (HIGH-) HIGHTECH RECEPTOR AB.  
XX  
XX Bjoerck L, Sjoerbring U;  
PI

XX  
DR WPI; 1993-368722/46.  
DR N-PSDB; AAQ50452.  
XX  
PT New protein L binding light chains of all immunoglobulin classes  
PT - for binding purifying and identifying immunoglobulin, also  
XX related DNA, vectors and host cells  
XX  
XX Claim 1; Page 36; 71pp; English.  
XX  
CC The protein (Protein L) is capable of binding to immunoglobulin G  
CC light chains. It is useful for binding, separating (purifying) and  
CC identifying immunoglobulin and for removing immunoglobulin molecules  
CC from serum. Hybrid proteins of the L protein can bind all human  
CC immunoglobulin classes and many immunoglobulins from other species.  
CC They are highly soluble and retain their binding activity at high  
CC temperatures over a pH range of 3-10. They can be immobilised  
CC without loss of activity.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 305 AA;  
Query Match 100.0%; Score 389; DB 14; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1.1e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPETPTDSEEEVTIKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKDNGEY 60  
Db 5 KEETPETPTDSEEEVTIKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKDNGEY 64  
QY 61 VDVAADKGYTLNKFAG 76  
Db 65 VDVAADKGYTLNKFAG 80  
RESULT 3  
AAR42994  
ID AAR42994 standard; Protein; 434 AA.  
XX AC AAR42994;  
XX  
XX 25-MAR-2003 (updated)  
DT 16-MAY-1994 (first entry)  
XX  
DE Sequence encoding immunoglobulin light chain binding protein.  
XX  
KW Immunoglobulin; light chain; binding; identification; purification;  
KW separation; ss.  
XX  
OS E. coli L392/pHDLG, DSM 7055.  
XX  
FH Key Location/Qualifiers  
FT Domain 5..305  
FT /label= B1 immunoglobulin light chain binding  
FT domain.  
FT Domain 81..305  
FT /label= B2 immunoglobulin light chain binding  
FT domain.  
FT Domain 153..305  
FT /label= B3 immunoglobulin light chain binding  
FT domain.  
FT Domain 225..305  
FT /label= B4 immunoglobulin light chain binding  
FT domain.  
FT Domain 297..305  
FT /label= B5 immunoglobulin light chain binding  
FT domain.  
FT Domain 309..434  
FT /label= C1 immunoglobulin heavy chain binding  
FT domain.  
FT Domain 364..434  
FT /label= D intermediate immunoglobulin heavy  
FT chain binding domain.  
FT

FT Domain 379..434  
 FT /label= C2 immunoglobulin heavy chain binding  
 FT domain.  
 XX WO9322342-A1.  
 XX 11-NOV-1993.  
 XX 28-APR-1993; 93WO-SE00375.  
 XX 28-APR-1992; 92SE-0001331.  
 XX (HIGH-) HIGTECH RECEPTOR AB.  
 XX Bjoerck L, Sjoerbring U;  
 XX WPI: 1993-368722/46.  
 XX P-PSDB; AAR42994.  
 XX New protein L binding light chains of all immunoglobulin classes  
 PT - for binding purifying and identifying immunoglobulin, also  
 PT related DNA, vectors and host cells  
 XX Claim 6; Page 39-40; 71pp; English.  
 XX Protein L (AAR42993) is capable of binding to immunoglobulin G light  
 CC chains. It is useful for binding, separating (purifying) and  
 CC identifying immunoglobulin and for removing immunoglobulin molecules  
 CC from serum. This is the coding sequence of one hybrid protein of  
 CC the L protein. The hybrid proteins can bind all human  
 CC immunoglobulin classes and many immunoglobulins from other species.  
 CC They are highly soluble and retain their binding activity at high  
 CC temperatures over a pH range of 3-10. They can be immobilised  
 CC without loss of activity.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 434 AA;  
 SQ

Query Match 100.0%; Score 389; DB 14; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-36;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KETPTPTDSEEVYTIKANLIFANGSTQTAEFGKTFEKATSEAYAYADTLKKNGEYT 60  
 DB 5 KETPTPTDSEEVYTIKANLIFANGSTQTAEFGKTFEKATSEAYAYADTLKKNGEYT 64  
 QY 61 VDVAADGYTLNIRFAG 76  
 DB 65 VDVAADGYTLNIRFAG 80

RESULT 4  
 AAB31372  
 ID AAB31372 standard; Protein; 467 AA.  
 XX AAB31372;  
 AC AAB31372;  
 XX 20-APR-2001 (first entry)  
 XX Amino acid sequence of protein L/CBD cex/ER retaining peptide fusion.  
 DE Protein production; food processing; protein antibiotic; feed enzyme;  
 XX protein L: CBD cex protein; cell signal peptide.  
 KW Synthetic.  
 XX WO20007174-A1.  
 XX 21-DEC-2000.  
 XX 07-JUN-2000; 2000WO-IL00330.  
 XX 10-JUN-1999; 99US-0329234.

XX (CBDT-) CBD TECHNOLOGIES LTD.  
 PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 XX Shani Z, Shoseyov O;  
 XX WPI: 2001-112219/12.  
 DR N-PSDB; AAF24730.  
 XX Expressing and isolating recombinant protein in a plant, useful for  
 PT producing large quantities of recombinant proteins, by expressing a  
 PT fusion protein including a cellulose binding peptide fused to a  
 PT recombinant protein  
 XX Example; Fig 2a; 87pp; English.  
 PS The specification describes a method for expressing and isolating a  
 CC recombinant protein in a plant. The method comprising expressing a  
 CC fusion protein including the recombinant protein and a cellulose  
 CC binding peptide fused to it, where the fusion protein is  
 CC compartmentalised and sequestered within plant cells, plant derived  
 CC tissue or cultured plant cells. The method is useful for obtaining large  
 CC quantities of the recombinant proteins and protein products in a simple  
 CC and cost-effective manner. Recombinant proteins may be used commercially,  
 CC such as in the food processing industry, e.g. glucosylases and glucose  
 CC isomerases are used for converting starch to high fructose corn syrup,  
 CC proteinases for the hydrolysis of high molecular weight proteins and in  
 CC manufacturing leather or alcoholic beverages, pectinesterases for  
 CC pectin hydrolysis in food industry, lipases for cleaving ester linkage  
 CC in triglycerides, and for effluent treatment. The recombinant proteins  
 CC may further be used to produce protein antibiotics, which can be used  
 CC in healing processes, and to produce animal feed enzymes. The present  
 CC sequence represents a fusion protein of the invention, and comprises a  
 CC fusion of a cell signal peptide, protein L, CBD cex, and an endoplasmic  
 CC reticulum retaining peptide.  
 XX Sequence 467 AA;  
 SQ

Query Match 98.7%; Score 384; DB 22; Length 467;  
 Best Local Similarity 98.7%; Pred. No. 7.4e-36;  
 Matches 75; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KETPTPTDSEEVYTIKANLIFANGSTQTAEFGKTFEKATSEAYAYADTLKKNGEYT 60  
 DB 44 KETPTPTDSEEVYTIKANLIFANGSTQTAEFGKTFEKATSEAYAYADTLKKNGEYT 103  
 QY 61 VDVAADGYTLNIRFAG 76  
 DB 104 VDVAADGYTLNIRFAG 119

RESULT 5  
 AAW32479  
 ID AAW32479 standard; Protein; 182 AA.  
 XX AAW32479;  
 AC AAW32479;  
 XX 27-MAR-1998 (first entry)  
 XX Growth factor LHL (catalytic antibody precursor).  
 DE Catalytic antibody; growth factor; B-cell mitogenesis; LHL;  
 XX L protein; hen egg lysozyme.  
 KW Chimeric - Peptostreptococcus magnus.  
 OS Chimeric - Gallus sp.  
 XX WO9735887-A1.  
 XX 02-OCT-1997.  
 XX 26-MAR-1997; 97WO-AU00194.  
 XX

```

PR 27-FEB-1997; 97AU-0005375.
PR 26-MAR-1996; 96AU-0008951.
XX
XX
XX (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
XX Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;
XX
XX WPI; 1997-489572/45.
XX DR N-PSDB; AAT91582.
XX
XX New catalytic antibody precursors - comprising a B-cell surface
XX molecule binding portion which can induce B-cell mitogenesis
XX
XX Claim 11; Page 56-57; 109pp; English.
XX
XX This polypeptide comprises growth factor LHL, where L is the
XX immunoglobulin binding entity from Peptostreptococcus magnus and H
XX is residues 42-62 of hen egg lysozyme. It was expressed in
XX Escherichia coli using a DNA construct (see AAT91586) generated by PCR.
XX Production of catalytic antibodies to a specific antigen comprises
XX administering to an animal a growth factor comprising an antigen
XX capable of interacting with a B cell bound catalytic antibody. The
XX antigen is fused to a B cell surface molecule binding protein for
XX the antigen to be cleaved and for the remainder of the molecule to
XX induce B cell mitogenesis (claimed). LHL crosslinks with surface
XX immunoglobulin on B cells. This induces B cell activation and blast
XX formation. Internalisation and processing of LHL leads to the
XX presentation of the H peptide on MHC II. T cell recognition of MHC
XX II with the H peptide signals the activated B cell to proliferate
XX and undergo antibody class switching and secretion. LHL has been
XX modified to contain a FLAG epitope and strep-tag (see AAW32481),
XX to contain a variable kappa light chain (see AAW32482) and to be
XX flanked by tumour necrosis factor (see AAW32480). The catalytic
XX antibodies generated by the process have diagnostic and therapeutic
XX applications.
XX
XX Sequence 182 AA;
XX
XX Query Match 83.8%; Score 326; DB 18; Length 182;
XX Best Local Similarity 97.0%; Pred. No. 1.1e-29;
XX Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 11 DSEEVTTKANLIFANGSTQTAETFGTPEKATSEAYAYADTLKDKNGEYTVADVADKGYTL 70
XX I: |||||
XX DB 25 DNTVEVTTKANLIFANGSTQTAETFGTPEKATSEAYAYADTLKDKNGEYTVADVADKGYTL 84
XX
XX QY 71 NIKFAG 76
XX |||||
XX DB 85 NIKFAG 90
XX
XX RESULT 6
XX AAY06907
XX ID AAY06907 standard; Protein; 182 AA.
XX AC
XX AAY06907;
XX
XX DT 01-JUL-1999 (first entry)
XX
XX LHL growth factor amino acid sequence.
XX
XX Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;
XX catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;
XX rheumatoid arthritis; toxic shock syndrome; viral docking receptor;
XX human immune virus; tumour-specific antigen; amyloid plaque; myeloma;
XX Alzheimer's disease; IGE; allergy; asthma; drug detoxification;
XX autoimmune; inflammatory disease; gene therapy; protein L; P. magnus;
XX hen egg lysozyme; HEL; LHL.
XX
XX Peptostreptococcus magnus.
XX Gallus sp.
XX
XX WO9915563-A1.

```

---

```

XX
XX PD 01-APR-1999.
XX
XX PF 18-SEP-1998; 98WO-AU00783.
XX
XX PR 19-SEP-1997; 97AU-0009306.
XX
XX PA (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
XX PI Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;
XX
XX DR WPI; 1999-244394/20.
XX DR N-PSDB; AAX34590.
XX
XX PT Growth factor precursor cleaved by antigen-specific catalytic
XX antibody
XX
XX Example 1; Page 62; 101pp; English.
XX
XX The invention relates to a growth factor precursor that comprises B-cell
XX surface binding part, T cell surface binding part, antigen cleavable
XX by a catalytic antibody (CAB); and a peptide comprising heavy and light
XX chains of immunoglobulin. When the antigen is cleaved the B cell surface
XX part can interact with its target. The growth factor precursors are used
XX to select B cells that produce Ag-specific CAB, and to generate CAB from
XX such cells (by inducing mitogenesis, caused by the growth factor released
XX by specific cleavage). The Ag-specific CAB can be directed against, e.g.
XX tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock
XX syndrome); viral docking receptors (treatment of human immune virus,
XX hepatitis and influenza infections); tumour-specific antigens; amyloid
XX plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of
XX allergies such as asthma). CAB may also be used for drug detoxification,
XX to treat autoimmune or inflammatory diseases and to eliminate
XX environmental or industrial pollutants, such as plastics and petroleum.
XX Particularly the growth factor precursors are produced by delivering the
XX corresponding nucleic acid in a viral or other gene therapy vector. The
XX present sequence represents the amino acid sequence of the LHL growth
XX factor. L is a B cell surface molecule binding portion from protein L of
XX P. magnus; H is a T cell surface molecule binding portion from hen egg
XX lysozyme (HEL). LHL is a growth factor comprising H flanked by two L
XX molecules.
XX
XX Sequence 182 AA;
XX
XX Query Match 83.8%; Score 326; DB 20; Length 182;
XX Best Local Similarity 97.0%; Pred. No. 1.1e-29;
XX Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 11 DSEEVTTKANLIFANGSTQTAETFGTPEKATSEAYAYADTLKDKNGEYTVADVADKGYTL 70
XX I: |||||
XX DB 25 DNTVEVTTKANLIFANGSTQTAETFGTPEKATSEAYAYADTLKDKNGEYTVADVADKGYTL 84
XX
XX QY 71 NIKFAG 76
XX |||||
XX DB 85 NIKFAG 90
XX
XX RESULT 7
XX AAW32486
XX ID AAW32486 standard; Protein; 178 AA.
XX AC
XX AAW32486;
XX
XX DT 27-MAR-1998 (first entry)
XX
XX DE Growth factor LHL-omp (catalytic antibody precursor).
XX
XX KW Catalytic antibody; growth factor; B-cell mitogenesis; LHL-omp;
XX L protein; hen egg lysozyme; HEL.
XX
XX OS Chimeric - Peptostreptococcus magnus.
XX OS Chimeric - Gallus sp.
XX OS Chimeric - Synthetic.

```

XX Key Location/Qualifiers  
 FH Peptide 2..9  
 FT /note= "FLAG epitope"  
 FT Peptide 170..178  
 FT /note= "strept-tag"  
 XX W09735887-AL.  
 XX 02-OCT-1997.  
 XX 26-MAR-1997; 97WO-AU00194.  
 XX 27-FEB-1997; 97AU-0005375.  
 XX 26-MAR-1996; 96AU-0008951.  
 XX (AMRA-) AMRAD OPERATIONS PTY LTD.  
 XX Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;  
 PI WPI; 1997-489572/45.  
 DR N-PSDB; AAT91591.  
 XX New catalytic antibody precursors - comprising a B-cell surface  
 PT molecule binding portion which can induce B-cell mitogenesis  
 XX  
 XX Example 27; Page 76; 109pp; English.  
 XX This polypeptide comprises growth factor LHL-omp, where L is the  
 CC immunoglobulin binding entity from Peptostreptococcus magnus and H  
 CC is residues 42-62 of hen egg lysozyme. It was generated from  
 CC LHL.seq (see AAW32481) such that the E. coli ompA signal sequence  
 CC (generated by purifying LHL.seq from whole bacterial host cells) was  
 CC eliminated. The N-terminal omp peptide found in LHL (see AAW32479),  
 CC LHL.seq (see AAW32481) and TLHL (see AAW32482) induces multimerisation  
 CC as demonstrated by a potentiation of biological activity as  
 CC compared to recombinant LHL-omp. Production of catalytic antibodies  
 CC to a specific antigen comprises administering to an animal a growth  
 CC factor comprising an antigen capable of interacting with a B cell  
 CC bound catalytic antibody. The antigen is fused to a B cell surface  
 CC molecule binding protein for the antigen to be cleaved and for the  
 CC remainder of the molecule to induce B cell mitogenesis (claimed).  
 CC The observation that omp induces multimerisation allows the design  
 CC of simpler molecules with the same desired biological function as  
 CC LHL, TLHL and CATAB (see AAW32480). The catalytic antibodies  
 CC generated by the process can have diagnostic and therapeutic  
 CC applications.  
 XX Sequence 178 AA;  
 SQ Query Match 83.0%; Score 323; DB 18; Length 178;  
 Best Local Similarity 98.5%; Pred. No. 2.3e-29;  
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 12 SEEEVIKANLIFANGSTQTAEPKGFKEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
 Db 102 SAEETVIKANLIFANGSTQTAEPKGFKEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161  
 QY 72 IKFAG 76  
 Db 162 IKFAG 166  
 RESULT 8  
 AAY06913  
 ID AAY06913 standard; Protein; 178 AA.  
 XX AAY06913;  
 AC  
 XX 01-JUL-1999 (first entry)  
 DT  
 XX Amino acid sequence of a form of LHL growth factor (LHL.omp).  
 DE  
 XX

KW Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;  
 KW catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;  
 KW rheumatoid arthritis; toxic shock syndrome; viral docking receptor;  
 KW human immune virus; tumour-specific antigen; amyloid plaque; myeloma;  
 KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification;  
 KW autoimmune; inflammatory disease; gene therapy; protein L; P. magnus;  
 KW hen egg lysozyme; HEL; LHL.  
 XX Peptostreptococcus magnus.  
 OS Gallus sp.  
 XX WO9915563-AL.  
 XX 01-APR-1999.  
 XX 18-SEP-1998; 98WO-AU00783.  
 XX 19-SEP-1997; 97AU-0009306.  
 XX (AMRA-) AMRAD OPERATIONS PTY LTD.  
 XX Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;  
 PI WPI; 1999-244394/20.  
 DR N-PSDB; AAX34595.  
 XX Growth factor precursor cleaved by antigen-specific catalytic  
 PT antibody  
 XX Examples; Page 74; 101pp; English.  
 XX The invention relates to a growth factor precursor that comprises B-cell  
 CC surface binding part, T cell surface binding part, antigen cleavable  
 CC by a catalytic antibody (CAB); and a peptide comprising heavy and light  
 CC chains of immunoglobulin. When the antigen is cleaved the B cell surface  
 CC part can interact with its target. The growth factor precursors are used  
 CC to select B cells that produce Ag-specific CAB, and to generate CAB from  
 CC such cells (by inducing mitogenesis, caused by the growth factor released  
 CC by specific cleavage). The Ag-specific CAB can be directed against, e.g.  
 CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock  
 CC syndrome); viral docking receptors (treatment of human immune virus,  
 CC hepatitis and influenza infections); tumour-specific antigens; amyloid  
 CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of  
 CC allergies such as asthma). CAB may also be used for drug detoxification,  
 CC to treat autoimmune or inflammatory diseases and to eliminate  
 CC environmental or industrial pollutants, such as plastics and petroleum.  
 CC Particularly the growth factor precursors are produced by delivering the  
 CC corresponding nucleic acid in a viral or other gene therapy vector. The  
 CC present sequence represents the amino acid sequence of the LHL growth  
 CC factor carrying an ompA signal peptide (LHL.omp). L is a B cell surface  
 CC molecule binding portion from protein L of P. magnus; H is a T cell  
 CC surface molecule binding portion from hen egg lysozyme (HEL). LHL is a  
 CC growth factor comprising H flanked by two L molecules.  
 XX Sequence 178 AA;  
 SQ Query Match 83.0%; Score 323; DB 20; Length 178;  
 Best Local Similarity 98.5%; Pred. No. 2.3e-29;  
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 12 SEEEVIKANLIFANGSTQTAEPKGFKEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
 Db 102 SAEETVIKANLIFANGSTQTAEPKGFKEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161  
 QY 72 IKFAG 76  
 Db 162 IKFAG 166  
 RESULT 9  
 AAW32481  
 ID AAW32481 standard; Protein; 198 AA.  
 XX







CC surface molecule binding portion from protein L of P. magnus; H is a T  
 CC cell surface molecule binding portion from hen egg lysozyme (HEL). TLHL  
 CC is LHL linked to kappa chain via TEV (tobacco etch virus) sensitive  
 CC peptide and g attached to N terminus region.  
 XX  
 SQ Sequence 342 AA;  
 Query Match 83.0%; Score 323; DB 20; Length 342;  
 Best Local Similarity 98.5%; Pred. No. 5.6e-29;  
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 12 SEEEVTKANLIFANGSTQTAEPKGFPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
 DB 174 SAEVTKANLIFANGSTQTAEPKGFPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 233  
 QY 72 IKFAG 76  
 DB 234 IKFAG 238  
 RESULT 13  
 AAY06915  
 ID AAY06915 standard; Protein; 482 AA.  
 AC AAY06915;  
 DT 01-JUL-1999 (first entry)  
 DE Amino acid sequence of a ccMTLgI growth factor.  
 KW Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;  
 KW catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;  
 KW rheumatoid arthritis; toxic shock syndrome; viral docking receptor;  
 KW human immune virus; tumour-specific antigen; amyloid plaque; myeloma;  
 KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification; ccMTLgI;  
 KW autoimmune; inflammatory disease; gene therapy; ccMTLgI; LGL.  
 OS  
 OS Peptostreptococcus magnus.  
 OS Synthetic.  
 OS  
 PN W09915563-A1. {  
 XX  
 PD 01-APR-1999.  
 XX  
 PF 18-SEP-1998; 98WO-AU00783.  
 XX  
 PR 19-SEP-1997; 97AU-0009306.  
 XX  
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.  
 XX  
 PI Koentgen F, Suess GM, Tarlington DM, Treutlein HR;  
 XX  
 DR WPI: 1999-244394/20.  
 DR N-PSDB; AAX34596.  
 XX  
 PT Growth factor precursor cleaved by antigen-specific catalytic  
 PT antibody  
 XX  
 PS Example 37; Page 77-78; 101pp; English.  
 XX  
 CC The invention relates to a growth factor precursor that comprises B-cell  
 CC surface binding part, T cell surface binding part, antigen cleavable  
 CC by a catalytic antibody (CAB); and a peptide comprising heavy and light  
 CC chains of immunoglobulin. When the antigen is cleaved the B cell surface  
 CC part can interact with its target. The growth factor precursors are used  
 CC to select B cells that produce Ag-specific CAB, and to generate CAB from  
 CC such cells (by inducing mitogenesis, caused by the growth factor released  
 CC by specific cleavage). The Ag-specific CAB can be directed against, e.g.  
 CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock  
 CC syndrome); viral docking receptors (treatment of human immune virus,  
 CC hepatitis and influenza infections); tumour-specific antigens; amyloid  
 CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of  
 CC allergies such as asthma). CAB may also be used for drug detoxification,

CC to treat autoimmune or inflammatory diseases and to eliminate  
 CC environmental or industrial pollutants, such as plastics and petroleum.  
 CC Particularly the growth factor precursors are produced by delivering the  
 CC corresponding nucleic acid in a viral or other gene therapy vector. The  
 CC present sequence represents the amino acid sequence of the ccMTLgI growth  
 CC factor. L is a B cell surface molecule binding portion from protein L of  
 CC P. magnus; LGL is two L molecules linked via glycine-serine peptide.  
 CC ccMTLgI is a growth factor comprising LgI linked to variable heavy and  
 CC light chain domains from antibody McpC603 via TEV (tobacco etch virus)  
 CC sensitive peptide.  
 XX  
 SQ Sequence 482 AA;  
 Query Match 83.0%; Score 323; DB 20; Length 482;  
 Best Local Similarity 98.5%; Pred. No. 8.9e-29;  
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 12 SEEEVTKANLIFANGSTQTAEPKGFPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
 DB 321 SAEVTKANLIFANGSTQTAEPKGFPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 380  
 QY 72 IKFAG 76  
 DB 381 IKFAG 385  
 RESULT 14  
 AAW32480  
 ID AAW32480 standard; Protein; 495 AA.  
 AC AAW32480;  
 XX  
 DT 27-MAR-1998 (first entry)  
 XX  
 DE Growth factor CATAB-TEV (catalytic antibody precursor).  
 XX  
 KW Catalytic antibody; growth factor; B-cell mitogenesis; LHL;  
 KW L protein; hen egg lysozyme; kappa light chain.  
 XX  
 OS Chimeric - Peptostreptococcus magnus.  
 OS Chimeric - Gallus sp.  
 OS Chimeric - Homo sapiens.  
 XX  
 PN W09735887-A1.  
 XX  
 PD 02-OCT-1997.  
 XX  
 PF 26-MAR-1997; 97WO-AU00194.  
 XX  
 PR 27-FEB-1997; 97AU-0005375.  
 XX  
 PR 26-MAR-1996; 96AU-0008951.  
 XX  
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.  
 XX  
 PI Koentgen F, Suess GM, Tarlington DM, Treutlein HR;  
 XX  
 DR WPI: 1997-489572/45.  
 DR N-PSDB; AAT91587.  
 XX  
 PT New catalytic antibody precursors - comprising a B-cell surface  
 PT molecule binding portion which can induce B-cell mitogenesis  
 XX  
 PS Claim 12; Page 61-63; 109pp; English.  
 XX  
 CC This polypeptide comprises CATAB-TEV, a synthetic growth factor  
 CC composed of: (a) TLHL (see AAW91589), where L is the immunoglobulin  
 CC binding entity of peptostreptococcus magnus, H is residues 42-62 of  
 CC hen egg lysozyme, and T represents the variable kappa light chain  
 CC (see AAW32483) from human myeloma protein LEN; and (b) an additional,  
 CC C-terminal kappa protein, the elements of CATAB-TEV being joined by  
 CC linkers containing recognition sites for tobacco etch virus (TEV)  
 CC protease. CATAB-TEV was designed so that the kappa portions of the  
 CC protein are removed by TEV protease to release 'T' and 'LHL'

CC moieties. Production of catalytic antibodies to a specific antigen  
 CC comprises administering to an animal a growth factor comprising an  
 CC antigen capable of interacting with a B cell bound catalytic  
 CC antibody. The antigen is fused to a B cell surface molecule  
 CC binding protein for the antigen to be cleaved and for the remainder  
 CC of the molecule to induce B cell mitogenesis (claimed). LHL (see  
 CC AA032479) crosslinks with surface immunoglobulin on B cells. This  
 CC induces B cell activation and blast formation. Internalisation  
 CC and processing of LHL leads to the presentation of the H peptide on  
 CC MHC II. T cell recognition of MHC II with the H peptide signals the  
 CC activated B cell to proliferate and undergo antibody class switching  
 CC and secretion. The catalytic antibodies generated by the process  
 CC can have diagnostic and therapeutic applications.

XX  
 SQ Sequence 495 AA;

Query Match 83.0%; Score 323; DB 18; Length 495;  
 Best Local Similarity 98.5%; Pred. No. 9.2e-29;  
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 12 SEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKDNGETVDVADKGYTLN 71  
 I|||||  
 Db 174 SAEVITKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKDNGETVDVADKGYTLN 233

OY 72 IKFAG 76  
 I|||||  
 Db 234 IKFAG 238

RESULT 15  
 AAY06908  
 ID AAY06908 standard; Protein; 495 AA.  
 AC AAY06908;  
 XX  
 XX 01-JUL-1999 (first entry)  
 DT  
 XX  
 DE CATAB-TEV aminoacid sequence.  
 XX  
 KW Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;  
 KW catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;  
 KW rheumatoid arthritis; toxic shock syndrome; viral docking receptor;  
 KW human immune virus; tumour-specific antigen; amyloid plaque; myeloma;  
 KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification;  
 KW autoimmune; inflammatory disease; gene therapy; protein L; P. magnus;  
 KW hen egg lysosyme; HEL; LHL; CATAB-TEV; tobacco etch virus; TLHL.  
 XX  
 OS Peptostreptococcus magnus.  
 OS Gallus sp.  
 XX  
 PN W09915563-A1.  
 PD 01-APR-1999.  
 XX  
 PF 18-SEP-1998; 98WO-AU00783.  
 XX  
 PR 19-SEP-1997; 97AU-0009306.  
 XX  
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.  
 XX  
 PI Koentgen F, Sues GM, Tarlington DM, Treutlein HR;  
 XX  
 DR WPI; 1999-244394/20.  
 DR N-PSDB; AAX34591.  
 XX  
 PT Growth factor precursor cleaved by antigen-specific catalytic  
 PT antibody  
 XX  
 PS Example 15; Page 64-66; 101pp; English.  
 XX  
 CC The invention relates to a growth factor precursor that comprises B-cell  
 CC surface binding part, T cell surface binding part, antigen cleavable  
 CC by a catalytic antibody (CAB); and a peptide comprising heavy and light

CC chains of immunoglobulin. When the antigen is cleaved the B cell surface  
 CC part can interact with its target. The growth factor precursors are used  
 CC to select B cells that produce Ag-specific CAB, and to generate CAB from  
 CC such cells (by inducing mitogenesis, caused by the growth factor released  
 CC by specific cleavage). The Ag-specific CAB can be directed against, e.g.  
 CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock  
 CC syndrome); viral docking receptors (treatment of human immune virus,  
 CC hepatitis and influenza infections); tumour-specific antigens; amyloid  
 CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of  
 CC allergies such as asthma). CAB may also be used for drug detoxification,  
 CC to treat autoimmune or inflammatory diseases and to eliminate  
 CC environmental or industrial pollutants, such as plastics and petroleum.  
 CC Particularly the growth factor precursors are produced by delivering the  
 CC corresponding nucleic acid in a viral or other gene therapy vector. The  
 CC present sequence represents the amino acid sequence of CATAB-TEV which  
 CC is assembled from TLHL and kappa by PCR. The LHL is a growth factor  
 CC comprising H flanked by two L molecules where L is a B cell surface  
 CC molecule binding portion from protein L of P. magnus; H is a T cell  
 CC surface molecule binding portion from hen egg lysosyme (HEL). TLHL is  
 CC LHL linked to kappa chain via TEV (tobacco etch virus) sensitive  
 CC peptide and g attached to N terminus region.

XX  
 SQ Sequence 495 AA;

Query Match 83.0%; Score 323; DB 20; Length 495;  
 Best Local Similarity 98.5%; Pred. No. 9.2e-29;  
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 12 SEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKDNGETVDVADKGYTLN 71  
 I|||||  
 Db 174 SAEVITKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKDNGETVDVADKGYTLN 233

OY 72 IKFAG 76  
 I|||||  
 Db 234 IKFAG 238

Search completed: September 3, 2003, 11:19:17  
 Job time : 52.0137 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:16:02 ; Search time 16.6575 Seconds  
(without alignments)  
193.043 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEETPTPTDSEEVTKA.....GEYTVDAKGYTLNIKFG 76

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	100.0	305	2	US-08-795-475-1
2	389	100.0	434	2	US-08-795-475-3
3	326	83.8	182	3	US-08-828-741B-2
4	326	83.8	182	4	US-09-160-567-2
5	326	83.8	182	4	US-09-710-299-2
6	323	83.0	178	3	US-08-828-741B-13
7	323	83.0	178	4	US-09-160-567-13
8	323	83.0	178	4	US-09-710-299-13
9	323	83.0	198	3	US-08-828-741B-8
10	323	83.0	198	4	US-09-160-567-8
11	323	83.0	198	4	US-09-710-299-8
12	323	83.0	342	3	US-08-828-741B-6
13	323	83.0	342	4	US-09-160-567-6
14	323	83.0	342	4	US-09-710-299-6
15	323	83.0	495	3	US-08-828-741B-4
16	323	83.0	495	4	US-09-160-567-4
17	323	83.0	495	4	US-09-710-299-4
18	298.5	76.7	1027	3	US-08-446-137B-2
19	298	76.6	71	3	US-08-446-137B-5
20	298	76.6	291	3	US-08-446-137B-4
21	284.5	73.1	75	3	US-08-446-137B-8
22	270.5	69.5	74	3	US-08-446-137B-7
23	262	67.4	71	3	US-08-446-137B-6
24	69	17.7	123	3	US-08-795-444-3
25	65	16.7	419	4	US-09-134-001C-3441
26	64.5	16.6	300	4	US-09-254-465A-10
27	63.5	16.3	8991	4	US-08-714-741-32

28	62.5	16.1	138	6	5218098-1	Patent No. 5218099
29	62.5	16.1	182	4	US-08-858-207A-434	Sequence 434, Appl
30	61.5	15.8	235	4	US-09-134-001C-4150	Sequence 4150, Ap
31	61	15.7	821	4	US-09-252-991A-21013	Sequence 21013, A
32	61	15.7	1094	4	US-09-268-347-32	Sequence 32, Appl
33	61	15.7	1167	2	US-08-589-756-2	Sequence 2, Appl
34	61	15.7	1167	3	US-09-206-800-2	Sequence 2, Appl
35	61	15.7	1167	4	US-09-206-898-2	Sequence 2, Appl
36	61	15.7	1181	4	US-09-206-898-23	Sequence 23, Appl
37	60	15.4	312	3	US-08-961-083-12	Sequence 12, Appl
38	60	15.4	312	4	US-09-536-784-12	Sequence 12, Appl
39	60	15.4	515	4	US-09-107-532A-5603	Sequence 5603, Ap
40	60	15.4	558	4	US-09-071-035-268	Sequence 268, App
41	60	15.4	1024	4	US-09-562-737-90	Sequence 90, Appl
42	60	15.4	1638	4	US-09-071-035-258	Sequence 258, App
43	60	15.4	1638	4	US-09-071-035-262	Sequence 262, App
44	60	15.4	1638	4	US-09-071-035-266	Sequence 266, App
45	60	15.4	3079	5	PCT-US94-00198-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-08-795-475-1  
; Sequence 1, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bjvreck, Lars  
; APPLICANT: Sjöbring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
US-08-795-475-1

Query Match 100.0%; Score 389; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 5.3e-41;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPTPTDSEEVTKANLIFANGSTGTAETKATSEAYAYADTLKDCGYT 60  
|||||

```
Db      5 KEETPETDSEEVTKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKKNGEYT 64
QY      61 VDADKGYTLNKFAG 76
        |||||
Db      65 VDADKGYTLNKFAG 80

RESULT 2
US-08-795-475-3
; Sequence 3, Application US/08795475
; Patent No. 5965390
; GENERAL INFORMATION:
; APPLICANT: Bjvrck, Lars
; APPLICANT: SJVdring, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,475
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.402D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055
US-08-795-475-3

Query Match      100.0%; Score 389; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.7e-41;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KEETPETDSEEVTKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKKNGEYT 60
        |||||
Db      5 KEETPETDSEEVTKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKKNGEYT 64

QY      61 VDADKGYTLNKFAG 76
        |||||
Db      65 VDADKGYTLNKFAG 80

RESULT 3
US-08-828-741B-2
; Sequence 2, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
```

```
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-828-741B-2

Query Match      83.8%; Score 326; DB 3; Length 182;
Best Local Similarity 97.0%; Pred. No. 2.3e-33;
Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      11 DSEEVTKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKKNGEYTVVDADKGYTL 70
        |||||
Db      25 DNTVEVTIKANLIFANGSTQTAEFGTPEKATSEAYAYADTLKKNGEYTVVDADKGYTL 84

QY      71 NIKFAG 76
        |||||
Db      85 NIKFAG 90

RESULT 4
US-09-160-567-2
; Sequence 2, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 182 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-2

Query Match 83.8%; Score 326; DB 4; Length 182;  
Best Local Similarity 97.0%; Pred. No. 2.3e-33;  
Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 11 DSEEVTKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKNNGEYTVVDVADKGYTL 70  
Db 25 DNTTEVTIKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKNNGEYTVVDVADKGYTL 84  
QY 71 NIKFAG 76  
Db 85 NIKFAG 90

RESULT 5  
US-09-710-299-2  
Sequence 2, Application US/09710299  
Patent No. 6521741  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
Suess, Gabriele M.  
Tarlinton, David M.  
Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/710,299  
FILING DATE: 09-No. 6521741-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 182 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-710-299-2

Query Match 83.8%; Score 326; DB 4; Length 182;  
Best Local Similarity 97.0%; Pred. No. 2.3e-33;  
Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 11 DSEEVTKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKNNGEYTVVDVADKGYTL 70  
Db 25 DNTTEVTIKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKNNGEYTVVDVADKGYTL 84  
QY 71 NIKFAG 76  
Db 85 NIKFAG 90

RESULT 6  
US-08-828-741B-13  
Sequence 13, Application US/08828741B  
Patent No. 6043069  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
Suess, Gabriele M.  
Tarlinton, David M.  
Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,741B  
FILING DATE: 26-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-828-741B-13

Query Match 83.0%; Score 323; DB 3; Length 178;  
Best Local Similarity 98.5%; Pred. No. 5.2e-33;  
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 12 SEEVTKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKNNGEYTVVDVADKGYTL 71  
I

Db 102 SAEVITKANLIFANGSTQTAETKGTFFKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161

QY 72 IKFAG 76  
|||||

Db 162 IKFAG 166

RESULT 7  
US-09-160-567-13  
; Sequence 13, Application US/09160567  
; Patent No. 6326179  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/160,567  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 178 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-160-567-13

Query Match 83.0%; Score 323; DB 4; Length 178;  
Best Local Similarity 98.5%; Pred. No. 5.2e-33;  
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 SAEVITKANLIFANGSTQTAETKGTFFKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
| |||||

Db 102 SAEVITKANLIFANGSTQTAETKGTFFKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161

QY 72 IKFAG 76  
|||||

Db 162 IKFAG 166

RESULT 8  
US-09-710-299-13  
; Sequence 13, Application US/09710299  
; Patent No. 6521741  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank

; Suess, Gabriele M.  
; Tarlinton, David M.  
; Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/710,299  
; FILING DATE: 09-No. 6521741-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 178 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-710-299-13

Query Match 83.0%; Score 323; DB 4; Length 178;  
Best Local Similarity 98.5%; Pred. No. 5.2e-33;  
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 SAEVITKANLIFANGSTQTAETKGTFFKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
| |||||

Db 102 SAEVITKANLIFANGSTQTAETKGTFFKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161

QY 72 IKFAG 76  
|||||

Db 162 IKFAG 166

RESULT 9  
US-08-828-741B-8  
; Sequence 8, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America





```

: Sequence 6, Application US/09160567
: Patent No. 6326179
: GENERAL INFORMATION:
: APPLICANT: Koentgen, Frank
: APPLICANT: Suess, Gabriele M.
: APPLICANT: Tarlinton, David M.
: APPLICANT: Treutlein, Herbert R.
: TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
: TITLE OF INVENTION: PRODUCING SAME
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
: STREET: 400 Garden City Plaza

```

CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-6

```

RESULT 14
US-09-710-299-6
; Sequence 6, Application US/09710299
; Patent No. 6521741
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; Suess, Gabriele M.
; Tarlinton, David M.
; Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
;

```

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza

RESULT 15  
 US-08-828-741B-4  
 ; Sequence 4, Application US/08828741B  
 ; Patent No. 6043069  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koentgen, Frank  
 ; APPLICANT: Suess, Gabriele M.  
 ; APPLICANT: Tarlinton, David M.  
 ; APPLICANT: Treutlein, Herbert R.  
 ; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
 ; TITLE OF INVENTION: PRODUCING SAME  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
 ; STREET: 400 Garden City plaza  
 ; CITY: Garden City  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 11530  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/828,741B  
 ; FILING DATE: 26-MAR-1997





APPLICANT: Bjvrck, Lars  
APPLICANT: Sjvbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 450023.401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
US-08-325-278-1

Query Match 100.0%; Score 389; DB 8; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1.6e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETDSEEEVTKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKNGEY 60  
|||||  
Db 5 KEETPETDSEEEVTKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKNGEY 64  
|||||

QY 61 VDVAADKGYTLNKFAG 76  
|||||  
Db 65 VDVAADKGYTLNKFAG 80  
|||||

RESULT 3  
US-08-325-278-3  
Sequence 3, Application US/08325278  
Publication No. US20030027283A1  
GENERAL INFORMATION:  
APPLICANT: Bjvrck, Lars  
APPLICANT: Sjvbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 450023.401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7055  
US-08-325-278-3

Query Match 100.0%; Score 389; DB 8; Length 434;  
Best Local Similarity 100.0%; Pred. No. 2.5e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETDSEEEVTKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKNGEY 60  
|||||  
Db 5 KEETPETDSEEEVTKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKNGEY 64  
|||||

QY 61 VDVAADKGYTLNKFAG 76  
|||||  
Db 65 VDVAADKGYTLNKFAG 80  
|||||

RESULT 4  
US-10-345-618-2  
Sequence 2, Application US/10345618  
Publication No. US20030148484A1  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
APPLICANT: Sues, Gabriele M.  
APPLICANT: Tarlinton, David R.  
APPLICANT: Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
FILE REFERENCE: 13474  
CURRENT APPLICATION NUMBER: US/10/345,618  
CURRENT FILING DATE: 2003-01-16  
PRIOR APPLICATION NUMBER: US/09/509,031  
PRIOR FILING DATE: 2000-06-09  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 182  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:LHL protein  
US-10-345-618-2

Query Match 83.8%; Score 326; DB 12; Length 182;  
Best Local Similarity 97.0%; Pred. No. 1.3e-29;  
Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 DSEEEVTKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKNGEYTVDAKGYTL 70  
|:|||||  
Db 25 DNTEEVTKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKNGEYTVDAKGYTL 84  
|:|||||

QY 71 NIKFAG 76  
|||||  
Db 85 NIKFAG 90  
|||||

Db 122 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 181

QY 72 IKFAG 76  
182 IKFAG 186

RESULT 7  
US-10-345-618-6  
; Sequence 6, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 342  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: TLHL protein  
; OTHER INFORMATION: sequence  
US-10-345-618-6

Query Match 83.0%; Score 323; DB 12; Length 342;  
Best Local Similarity 98.5%; Pred. No. 6.5e-29;  
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
174 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 233

QY 72 IKFAG 76  
234 IKFAG 238

RESULT 8  
US-10-345-618-16  
; Sequence 16, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: ccmTlgL protein  
; OTHER INFORMATION: sequence  
US-10-345-618-16

Query Match 83.0%; Score 323; DB 12; Length 482;

RESULT 5  
US-10-345-618-13  
; Sequence 13, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: LHL-omp  
; OTHER INFORMATION: protein sequence  
US-10-345-618-13

Query Match 83.0%; Score 323; DB 12; Length 178;  
Best Local Similarity 98.5%; Pred. No. 2.8e-29;  
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
102 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161

QY 72 IKFAG 76  
162 IKFAG 166

RESULT 6  
US-10-345-618-8  
; Sequence 8, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: LHL.seq  
; OTHER INFORMATION: protein sequence  
US-10-345-618-8

Query Match 83.0%; Score 323; DB 12; Length 198;  
Best Local Similarity 98.5%; Pred. No. 3.2e-29;  
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 71  
102 SAEVITKANLIFANGSTQAEFGKTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 161

Query Match	80.7%	Score 314;	DB 10;	Length 72;
Best Local Similarity	81.6%	Pred. No. 9.3e-29;		

```

Query Match      76.6%; Score 298; DB 10; Length 71;
Best Local Similarity 85.9%; Pred. No. 6.le-27;
Matches 61; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 6 ETPTOSEEVTKANLIFANGSTQTAEPFGTPEKATSEAYAYADTLKKDNGEYTVDDVAD 65
      |||||-----|||-----|||-----|||-----|||-----|||-----|||
Db 2 ETPE--PEEVTTKANLIFADGSTQNAEPFGTFAKAVSDAYAYADALKKDNGEYTVDDVAD 59
      |||||-----|||-----|||-----|||-----|||-----|||-----|||

QY 66 KGVTLNLIKFAAG 76

```



```
Db          60 KGLTLNKFAG 70
          || |||||
RESULT 13
US-09-808-212A-18
; Sequence 16, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-18
Query Match          73.1%; Score 284.5; DB 10; Length 75;
Best Local Similarity 74.7%; Pred. No. 2.3e-25;
Matches 56; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

QY 2 EETPETPDSEEEVTKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTV 61
      :||||| : ||||| ||||| : ||||| ||||| : ||||| ||||| : |||||
Db 1 KETPETPE-EPKEEVIKYNLIFADGKTQTAEFKGTPEEATAEAYAYADLLAKVNGEYTA 59
      : |||||
QY 62 DVADKGYTLNKFAG 76
      | : |||||
Db 60 DLEDGGYTNIRFAG 74

RESULT 14
US-09-808-212A-16
; Sequence 16, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-16
Query Match          69.5%; Score 270.5; DB 10; Length 74;
Best Local Similarity 70.7%; Pred. No. 9e-24;
Matches 53; Conservative 11; Mismatches 10; Indels 1; Gaps 1;

QY 2 EETPETPDSEEEVTKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTV 61
      :||||| : ||||| ||||| : ||||| ||||| : ||||| ||||| : |||||
Db 1 KETPETPE-EPKEEVIKYNLIFADGKTQTAEFKGTPEEATAEAYAYANLLAKENG EYTA 59
      : |||||
QY 62 DVADKGYTLNKFAG 76
      | : |||||
Db 60 DLEDGGYTNIRFAG 74

Search completed: September 3, 2003, 11:42:50
Job time : 29.1096 secs
```

```
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-10
Query Match          69.2%; Score 269; DB 10; Length 72;
Best Local Similarity 68.4%; Pred. No. 1.3e-23;
Matches 52; Conservative 11; Mismatches 9; Indels 4; Gaps 1;

QY 1 KEETPETPDSEEEVTKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTV 60
      ||||| | : ||||| ||||| : ||||| ||||| : ||||| ||||| : |||||
Db 1 KEKTPPEP----KEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYAYADLLAKENGKYT 56
      | : |||||
QY 61 VDADKGYTLNKFAG 76
      | : |||||
Db 57 ADLEDGGYTNIRFAG 72

Search completed: September 3, 2003, 11:42:50
Job time : 29.1096 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:19:27 ; Search time 224.877 Seconds  
(without alignments)  
294.149 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEETPETPDSSEEVTIKAKLIFANGSTOTAFKGTFFKATSEYAYATLKDNGEYT 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

```

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	100.0	76	23	US-09-808-212A-4 Sequence 4, Appli
2	389	100.0	78	22	US-09-791-537-22553 Sequence 22553, A

3	389	100.0	305	7	US-08-325-278-1	Sequence 1, Appli
4	389	100.0	305	7	US-08-325-278A-1	Sequence 1, Appli
5	389	100.0	305	7	US-08-325-278B-1	Sequence 1, Appli
6	389	100.0	434	7	US-08-325-278-3	Sequence 3, Appli
7	389	100.0	434	7	US-08-325-278A-3	Sequence 3, Appli
8	389	100.0	434	7	US-08-325-278B-3	Sequence 3, Appli
9	389	100.0	719	22	US-09-791-537-10210	Sequence 10210, A
10	389	100.0	719	22	US-09-791-537-96101	Sequence 96101, A
11	384	98.7	467	25	US-09-980-469-12	Sequence 12, Appli
12	326	83.8	182	19	US-09-509-031-2	Sequence 2, Appli
13	326	83.8	182	23	US-09-820-048A-2	Sequence 2, Appli
14	323	83.0	178	19	US-09-509-031-13	Sequence 13, Appli
15	323	83.0	178	23	US-09-820-048A-13	Sequence 13, Appli
16	323	83.0	198	23	US-09-509-031-8	Sequence 8, Appli
17	323	83.0	198	23	US-09-820-048A-8	Sequence 8, Appli
18	323	83.0	342	19	US-09-509-031-6	Sequence 6, Appli
19	323	83.0	342	23	US-09-820-048A-6	Sequence 6, Appli
20	323	83.0	482	19	US-09-509-031-16	Sequence 16, Appli
21	323	83.0	495	19	US-09-509-031-4	Sequence 4, Appli
22	323	83.0	495	23	US-09-820-048A-4	Sequence 4, Appli
23	314	80.7	72	23	US-09-808-212A-6	Sequence 6, Appli
24	314	80.7	367	23	US-09-889-182A-4	Sequence 4, Appli
25	299	76.9	72	23	US-09-808-212A-8	Sequence 8, Appli
26	298.5	76.7	992	22	US-09-791-537-88366	Sequence 88366, A
27	298.5	76.7	1027	7	US-08-331-637-2	Sequence 2, Appli
28	298.5	76.7	1027	8	US-08-446-137A-2	Sequence 2, Appli
29	298.5	76.7	1027	15	US-09-187-295-2	Sequence 2, Appli
30	298	76.6	71	23	US-09-808-212A-12	Sequence 12, Appli
31	298	76.6	71	8	US-08-446-137A-4	Sequence 4, Appli
32	284.5	73.1	75	23	US-09-808-212A-18	Sequence 18, Appli
33	270.5	69.5	74	23	US-09-808-212A-16	Sequence 16, Appli
34	269	69.2	72	23	US-09-808-212A-10	Sequence 10, Appli
35	262.5	67.5	82	23	US-09-808-212A-2	Sequence 2, Appli
36	262	67.4	71	23	US-09-808-212A-14	Sequence 14, Appli
37	74	19.0	269	19	US-09-540-2098-7625	Sequence 7625, Ap
38	74	19.0	529	22	US-09-791-537-131077	Sequence 131077, A
39	73	18.8	634	22	US-09-791-537-126395	Sequence 126395, A
40	72	18.5	243	21	US-09-708-427-12618	Sequence 12618, A
41	72	18.5	487	21	US-09-708-427-12617	Sequence 12617, A
42	72	18.5	487	22	US-09-791-537-28306	Sequence 28306, A
43	71.5	18.4	841	1	PCT-US02-03987-5779	Sequence 5779, A
44	71.5	18.4	841	1	PCT-US02-03987-12751	Sequence 12751, A
45	71.5	18.4	841	23	US-09-815-242-5779	Sequence 5779, Ap

#### ALIGNMENTS

RESULT 1  
US-09-808-212A-4  
; Sequence 4, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Iman Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-4

Query Match 100.0%; Score 389; DB 23; Length 76;  
Best Local Similarity 100.0%; Pred. No. 3.6e-37;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETPDSSEEVTIKAKLIFANGSTOTAFKGTFFKATSEYAYATLKDNGEYT 60  
|||||





REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 100084.402  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-08-325-278A-3  
Query Match 100.0%; Score 389; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 3.9e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 60  
Db 5 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 64  
QY 61 VDVAADKGYTLNKFAG 76  
Db 65 VDVAADKGYTLNKFAG 80  
RESULT 8  
US-08-325-278B-3  
; Sequence 3, Application US/08325278B  
; GENERAL INFORMATION:  
; APPLICANT: Bjorck, Lars  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278B  
; FILING DATE: 26-Oct-1994  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E. R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 100084.402  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-08-325-278B-3  
Query Match 100.0%; Score 389; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 3.9e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 60  
Db 5 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 64  
QY 61 VDVAADKGYTLNKFAG 76  
Db 65 VDVAADKGYTLNKFAG 80  
RESULT 9  
US-09-791-537-10210  
; Sequence 10210, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 10210  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus magnus  
US-09-791-537-10210  
Query Match 100.0%; Score 389; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 7.7e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 60  
Db 98 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 157  
QY 61 VDVAADKGYTLNKFAG 76  
Db 158 VDVAADKGYTLNKFAG 173  
RESULT 10  
US-09-791-537-96101  
; Sequence 96101, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 96101  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus magnus  
US-09-791-537-96101  
Query Match 100.0%; Score 389; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 7.7e-36;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEETPETDSEEEVTIKANLIFANGSTQTAFFKGTFEKATSEAYAYADTLKKDNGEY 60

Db 98 KEETPETDSEEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTL 157  
QY 61 VDADKGYTLNFKAG 76  
Db 158 VDADKGYTLNFKAG 173

RESULT 11  
US-09-980-469-12  
; Sequence 12, Application US/09980469  
; GENERAL INFORMATION:  
; APPLICANT: Shoseyov, Oded  
; APPLICANT: Ziv, Shani  
; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND RECO  
; FILE OF INVENTION: PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTURED  
; FILE REFERENCE: 01/22924  
; CURRENT APPLICATION NUMBER: US/09/980,469  
; CURRENT FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: puc19-cell-Protl-cexNG-HDEL fusion encoded product  
US-09-980-469-12

Query Match 98.7%; Score 384; DB 25; Length 467;  
Best Local Similarity 98.7%; Pred. No. 1.6e-35;  
Matches 75; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KEETPETDSEEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTL 60  
Db 44 KEETPETDSEEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTL 103  
QY 61 VDADKGYTLNFKAG 76  
Db 104 VDADKGYTLNFKAG 119

RESULT 12  
US-09-509-031-2  
; Sequence 2, Application US/09509031  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/09/509,031  
; CURRENT FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:LHL protein  
; OTHER INFORMATION: sequence  
US-09-509-031-2

Query Match 83.8%; Score 326; DB 19; Length 182;  
Best Local Similarity 97.0%; Pred. No. 2.6e-29;  
Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 11 DSEEEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTVADKGYTL 70  
Db 25 DNTTEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTVADKGYTL 84

QY 71 NIKFAG 76  
Db 85 NIKFAG 90

RESULT 13  
US-09-820-048A-2  
; Sequence 2, Application US/09820048A  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/820,048A  
; FILING DATE: 28-Mar-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-820-048A-2

Query Match 83.8%; Score 326; DB 23; Length 182;  
Best Local Similarity 97.0%; Pred. No. 2.6e-29;  
Matches 64; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 DSEEEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTVADKGYTL 70  
Db 25 DNTTEVTKANLIFANGSTQTAFFKGTSEAYAYADTLKDKNGEYTVADKGYTL 84

QY 71 NIKFAG 76  
Db 85 NIKFAG 90

RESULT 14  
US-09-509-031-13  
; Sequence 13, Application US/09509031  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.

US-09-820-048A-13

Query Match 83.0%; Score 323; DB 23; Length 178;  
Best Local Similarity 98.5%; Pred. No. 5.7e-29;  
Matches 64; Conservative 0; Mismatches 1; Indels

Search completed: September 3, 2003, 11:40:07  
Job time : 225.877 secs

## RESULT 15

US-09-820-048A-13  
: Sequence 13: Application US/09820048A

GENERAL INFORMATION:  
APPLICANT: Koentgen Frank

Suess, Gabriele M.

;  
;  
Treutlein, Herbert

PRODUCING

NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SCULLY, SCOTT,

STREET: 400 Garden City P  
CITY: Garden City

STATE: New York  
COUNTRY: United states of

ZIP: 11530

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS 3.31

```
, SOFTWARE: PATENTLII RELEASES  
:  
: CURRENT APPLICATION DATA:
```

FILING DATE: 28-Mar-2001

;  
;  
PRIOR APPLICATION DATA:

FILING DATE: <Unknown>

NAME: DIGIGLIO, Frank S.  
REGISTRATION NUMBER: 21 3

REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 13:

LENGTH: 179 amino acids

TOPOLOGY: linear

```

;
SEQUENCE DESCRIPTION: SEQ ID NUMBER

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:20:12 ; Search time 10.6712 Seconds  
(without alignments)  
223.701 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEEPTPETDSEEVTKA.....GEYTVADVADKGYTLNIKPAQ 76

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 148256 seqs, 31410045 residues

Total number of hits satisfying chosen parameters: 148256

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pcp.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71.5	18.4	930	5	US-09-200-650E-3
2	64.5	16.6	102	6	US-10-016-516A-12
3	64.5	16.6	300	5	US-09-397-243D-12
4	64.5	16.6	300	6	US-10-633-008-10
5	63.5	16.3	170	6	US-10-299-636-47
6	61.5	15.8	284	6	US-10-326-956-2287
7	61.5	15.8	1290	6	US-10-617-320-4399
8	61.5	15.8	1303	6	US-10-640-833-5037
9	61.5	15.8	1372	6	US-10-273-573-8150
10	61.5	15.8	1377	6	US-10-273-573-8152
11	61	15.7	253	5	US-09-897-516A-7245
12	61	15.7	311	6	US-10-617-320-2862
13	61	15.7	311	6	US-10-640-833-4119
14	60.5	15.6	130	6	US-10-299-636-32
15	60.5	15.6	271	6	US-10-221-131-95
16	60.5	15.6	271	6	US-10-221-131-96
17	60.5	15.6	1166	5	US-09-200-650E-7
18	60	15.4	283	6	US-10-603-113-15180
19	60	15.4	301	6	US-10-617-320-4522
20	60	15.4	335	6	US-10-640-833-4355
21	60	15.4	342	6	US-10-617-320-3903
22	60	15.4	399	6	US-10-640-833-2754
23	60	15.4	1302	6	US-10-322-696-57
24	60	15.4	2602	1	PCT-US03-21510-96
25	59.5	15.3	239	6	US-10-603-113-15852
26	59.5	15.3	472	6	US-10-286-897-2428

27 59.5 15.3 472 6 US-10-258-898A-2428 Sequence 2428, Ap  
28 59.5 15.3 472 7 US-60-485-450-979 Sequence 979, Ap  
29 59.5 15.3 485 6 US-10-286-897-6000 Sequence 6000, Ap  
30 59.5 15.3 485 6 US-10-258-898A-6000 Sequence 6000, Ap  
31 59.5 15.3 609 6 US-10-603-113-20277 Sequence 20277, A  
32 59.5 15.3 1313 7 US-60-478-196-3167 Sequence 3167, Ap  
33 59 15.2 357 1 PCT-US02-14753A-826 Sequence 826, Ap  
34 59 15.2 484 6 US-10-603-113-14385 Sequence 14385, A  
35 59 15.2 692 6 US-10-326-956-1061 Sequence 1061, Ap  
36 59 15.2 753 6 US-10-603-114-5022 Sequence 5022, Ap  
37 59 15.2 1228 1 PCT-US02-14753A-537 Sequence 537, Ap  
38 59 15.2 1261 1 PCT-US02-14753A-538 Sequence 538, Ap  
39 59 15.2 1261 1 PCT-US02-14753A-1009 Sequence 1009, Ap  
40 59 15.2 1303 6 US-10-322-696-60 Sequence 60, Appl  
41 59 15.2 1316 6 US-10-273-573-10242 Sequence 10242, A  
42 59 15.2 1464 1 PCT-US02-14753A-1008 Sequence 1008, Ap  
43 58.5 15.0 464 6 US-10-275-595A-14 Sequence 14, Appl  
44 58.5 15.0 464 6 US-10-293-244-1211 Sequence 1211, Ap  
45 58.5 15.0 606 6 US-10-603-113-14276 Sequence 14276, A

#### ALIGNMENTS

##### RESULT 1

US-09-200-650E-3

; Sequence 3, Application US/09200650E

; GENERAL INFORMATION:

; APPLICANT: Patti, Joseph M.

; APPLICANT: Foster, Timothy J.

; APPLICANT: Hook, Magnus A.O.

; APPLICANT: Eihinn, Deirdre N1

; APPLICANT: Perkins, Samuel L.

; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aure

; FILE REFERENCE: P06283052/BAS

; CURRENT APPLICATION NUMBER: US/09/200,650E

; CURRENT FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: 60/066,815

; PRIOR FILING DATE: 1997-11-26

; PRIOR APPLICATION NUMBER: 60/098,427

; PRIOR FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 930

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-200-650E-3

Query Match 18.4%; Score 71.5; DB 5; Length 930;  
Best Local Similarity 30.6%; Pred. No. 6.4;  
Matches 22; Conservative 17; Mismatches 22; Indels 11; Gaps 5;

QY 7 TPETSESEVTKANLIFAN-GSTOTAE-FKTFKATSEAY-----AYADTLKKDNG-- 57

DB 407 TPDTSKLDVTDQDFVIYSNDKNTATVLMKG--QTSSNKQYIIQQVAYPDNSSTDNGKI 464

QY 58 EYTVDVADKGYT 69

DB 465 DYTLDTKTKYS 476

##### RESULT 2

US-10-016-516A-12

; Sequence 12, Application US/10016516A

; GENERAL INFORMATION:

; APPLICANT: Houtzager, Erwin

; APPLICANT: Vijjn, Irma Maria Caecilia

; APPLICANT: Sijmons, Peter Christiaan

; TITLE OF INVENTION: A Structure for Presenting Desired Peptide Sequences

; FILE REFERENCE: 2183-5208US

; CURRENT APPLICATION NUMBER: US/10/016,516A

; CURRENT FILING DATE: 2002-05-21

```

; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/10/265,542
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US/09/254,465
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US99/05028
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-633-008-10

Query Match          16.6%; Score 64.5; DB 6; Length 300;
Best Local Similarity 26.2%; Pred. No. 9.4;
Matches 22; Conservative 14; Mismatches 19; Indels 29; Gaps 4;

QY      6   ETPETDSEEVTKANLIIFANGSTQTAEFGTTEKATSEAYA-----YADTL----- 52
Db       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
38 QVPENES----IKLCTVSGFSSPRVEWK--FVGSTALVCYNISOITAPYADVTFESS 90
           : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
           : ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      53   -----KKNGEYTVDVADKG 67
Db       : ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
91 SGITFSVTNRKDNGEYTCMVSEEG 114

RESULT 5
US-10-299-636-47
; Sequence 47, Application US/102999636
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yoether, Janet
; APPLICANT: Crain, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (167)
; OTHER INFORMATION: Xaa at position 167 is unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (169)
; OTHER INFORMATION: Xaa at position 169 is unknown
US-10-299-636-47

Query Match          16.3%; Score 63.5; DB 6; Length 170;
Best Local Similarity 45.0%; Pred. No. 6;
Matches 18; Conservative 3; Mismatches 18; Indels 1; Gaps 1;

```

REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-011

US-10-640-833-5037

117 KAEASPKEEAKEVDSKESNTDKTDK 143

```
RESULT 9
US-10-273-573-8150
; Sequence 8150, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-086
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8150
; LENGTH: 1372
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (795)..(810)
; OTHER INFORMATION: WILM'S TUMOUR PROTEIN SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00049D, p-value=7.500e-10, raw score of 0.00
US-10-273-573-8150

Query Match 15.8%; Score 61.5; DB 6; Length 1372;
Best Local Similarity 24.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 12; Mismatches 26; Indels 21; Gaps 2;

QY 1 KEETPTPTD-----SEEVTKANLIFANGSTOTAEKGTFEKA 41
Db 35 EEEPRLPSSDLGGVPWKEAVRIHALLKKGSEEE--LEASKSPGPGNEEEEEEEYEE 92

QY 42 TSEAYAYADTLKKDNGEY 59
Db 93 EEDYDEEEESSEGEY 110

RESULT 10
US-10-273-573-8152
; Sequence 8152, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-086
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8152
; LENGTH: 1377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (795)..(810)
; OTHER INFORMATION: WILM'S TUMOUR PROTEIN SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00049D, p-value=7.500e-10, raw score of 0.00
US-10-273-573-8152

Query Match 15.8%; Score 61.5; DB 6; Length 1377;
Best Local Similarity 24.4%; Pred. No. 1.3e+02;
Matches 19; Conservative 12; Mismatches 26; Indels 21; Gaps 2;

QY 1 KEETPTPTD-----SEEVTKANLIFANGSTOTAEKGTFEKA 41
Db 35 EEEPRLPSSDLGGVPWKEAVRIHALLKKGSEEE--LEASKSPGPGNEEEEEEEYEE 92
```

```
QY 42 TSEAYAYADTLKKDNGEY 59
Db 93 EEDYDEEEESSEGEY 110

RESULT 11
US-09-897-516A-7245
; Sequence 7245, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 7245
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-7245

Query Match 15.7%; Score 61; DB 5; Length 253;
Best Local Similarity 31.3%; Pred. No. 18;
Matches 26; Conservative 9; Mismatches 18; Indels 30; Gaps 6;

QY 9 ETDSEEV-----IKANLIFANGSTOTAEKGTFEKATSEAYAYADTLKKDNGEY 60
Db 53 DTDSDSEIVATSKSVKAYDLANSANQNA-----SSANDNATLSKDRNG---- 99

QY 61 VDVADK-----GYTLNKF 74
Db 100 ADIPDKTNELKNIGLGETANVR 122

RESULT 12
US-10-617-320-2862
; Sequence 2862, Application US/10617320
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/617,320
; FILING DATE: 10-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
```

FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 2862:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 311 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...311  
SEQUENCE DESCRIPTION: SEQ ID NO: 2862:  
US-10-617-320-2862

Query Match 15.7%; Score 61; DB 6; Length 311;  
Best Local Similarity 30.0%; Pred. No. 23;  
Matches 18; Conservative 7; Mismatches 33; Indels 2; Gaps 1;

QY 10 TDSSEVITKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKDNGEYTVADVADKGYT 69  
Db 112 TDSIEIVNPDHYLFTIG--EGSSLKATMTVNSGRGVVPADENKKDNAPVGTGLAVDSIYT 169

## RESULT 13

US-10-640-833-4119  
Sequence 4119, Application US/10640833  
GENERAL INFORMATION:  
APPLICANT: Doucette-Stamm, Lynn  
APPLICANT: Bush, David  
APPLICANT: Zeng, Qiangdong  
APPLICANT: Opperman, Timothy  
APPLICANT: Houseweart, Chad Eric  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
FILE REFERENCE: PATH03-15  
CURRENT APPLICATION NUMBER: US/10/640,833  
CURRENT FILING DATE: 2003-08-14  
PRIOR APPLICATION NUMBER: US 09/583,110  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107,433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085,131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051,553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 4119  
LENGTH: 311  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-10-640-833-4119

Query Match 15.7%; Score 61; DB 6; Length 311;  
Best Local Similarity 30.0%; Pred. No. 23;  
Matches 18; Conservative 7; Mismatches 33; Indels 2; Gaps 1;

QY 10 TDSSEVITKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKDNGEYTVADVADKGYT 69  
Db 112 TDSIEIVNPDHYLFTIG--EGSSLKATMTVNSGRGVVPADENKKDNAPVGTGLAVDSIYT 169

## RESULT 14

US-10-299-636-32  
Sequence 32, Application US/10299636

GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 32  
LENGTH: 130  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-10-299-636-32

Query Match 15.6%; Score 60.5; DB 6; Length 130;  
Best Local Similarity 50.0%; Pred. No. 9.3;  
Matches 17; Conservative 2; Mismatches 14; Indels 1; Gaps 1;

QY 22 LIFANGSTQTAEFKGTFEKATSEAYAYADTLKKD 55  
Db 28 LVRAEESPOVVE--KSSLEKRYEAKAKADTAKKD 60

## RESULT 15

US-10-221-131-95  
Sequence 95, Application US/10221131  
GENERAL INFORMATION:  
APPLICANT: FUKUSHIMA, NAOSHI  
APPLICANT: TSUCHIYA, MASAYUKI  
APPLICANT: OH-EDA, MASAYOSHI  
APPLICANT: UNO, SHINSUKE  
APPLICANT: KIKUCHI, YASUFUMI  
TITLE OF INVENTION: POLYPEPTIDE INDUCING APOPTOSIS  
FILE REFERENCE: 065678/0106  
CURRENT APPLICATION NUMBER: US/10/221,131  
CURRENT FILING DATE: 2003-02-11  
PRIOR APPLICATION NUMBER: PCT/JP01/01912  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: 09/523,095  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: JP 2000-115246  
PRIOR FILING DATE: 2000-04-17  
PRIOR APPLICATION NUMBER: JP 2000-321822  
PRIOR FILING DATE: 2000-10-20  
NUMBER OF SEQ ID NOS: 103  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 95  
LENGTH: 271  
TYPE: PRT  
ORGANISM: Mus sp.  
FEATURE:  
OTHER INFORMATION: pscM1. MABL1-scfv  
US-10-221-131-95

Query Match 15.6%; Score 60.5; DB 6; Length 271;  
Best Local Similarity 27.3%; Pred. No. 23;  
Matches 18; Conservative 12; Mismatches 23; Indels 13; Gaps 2;

QY 24 FANGSTQTAEFKGTFEKATSEAYAYADTLKKD-----GEYTVADVADKGYTL 70  
Db 76 YNDGTYNEKFKGKATUTSEKSSAAYNELSSLASDSAVYCYARGGYYSYDDWGQCTTL 135

QY 71 NIKFAG 76  
Db 136 TVSSGG 141

Search completed: September 3, 2003, 11:40:55  
Job time : 11.6712 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:14:37 ; Search time 17.9589 Seconds  
(without alignments)  
406.975 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEETPETDSEEEVTKA.....GEYTVADKGYTLNIKFPAG 76

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	389	100.0	455	2 A45063	immunoglobulin-bi
2	389	100.0	719	2 A42808	Ig light chain-bin
3	298.5	76.7	992	2 S54396	protein L precursor
4	144.5	37.1	74	2 A34483	Ig light chain-bin
5	85	21.9	1576	2 A80249	probable hemolysin
6	74	19.0	529	2 S62194	hypothetical prote
7	72	18.5	487	2 T10215	hypothetical prote
8	71	18.3	528	2 D83874	nickel ABC transpo
9	71	18.3	929	2 T51932	kinesin [imported]
10	69.5	17.9	292	2 AGL457	p45 related protei
11	69	17.7	123	2 T80601	hypothetical prote
12	69	17.7	341	2 S28641	hypothetical prote
13	67.5	17.4	1056	2 B82557	hypothetical prote
14	67.5	17.4	1876	2 E97944	zinc metalloprotei
15	67	17.2	571	2 A43507	streptolysin O pre
16	67	17.2	779	2 T05990	hypothetical prote
17	67	17.2	890	2 S22452	surface exclusion
18	67	17.2	891	2 G41662	130K surface exclu
19	66.5	17.1	255	2 F89769	hypothetical prote
20	66.5	17.1	642	2 H81868	probable virulence
21	66.5	17.1	840	2 S48975	hypothetical prote
22	66.5	17.1	843	2 S72375	surface exclusion
23	66	17.0	183	2 E83809	hypothetical prote
24	66	17.0	498	2 A86014	cell division memb
25	66	17.0	498	2 A91168	cell division memb
26	66	17.0	653	2 F70383	organic solvent to
27	65.5	16.8	313	2 T33391	hypothetical prote
28	64.5	16.6	273	2 T47891	hypothetical prote
29	64	16.5	298	2 T47760	hypothetical prote

30 64 16.5 299 2 T33390  
31 64 16.5 316 2 G71600  
32 64 16.5 478 2 A71758  
33 64 16.5 582 2 S42613  
34 64 16.5 583 2 F69153  
35 63.5 16.3 442 1 IEEC41  
36 63.5 16.3 568 2 E97066  
37 63.5 16.3 744 2 F95013  
38 63 16.2 285 2 C89959  
39 63 16.2 636 2 H70184  
40 63 16.2 799 2 T02456  
41 63 16.2 815 2 T46169  
42 63 16.2 954 2 I51703  
43 63 16.2 3063 2 S55505  
44 63 16.2 3562 2 A47171  
45 62.5 16.1 280 2 I55577

## ALIGNMENTS

### RESULT 1

A45063

immunoglobulin-binding protein LG - Peptostreptococcus magnus

C:Species: Peptostreptococcus magnus

C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999

C:Accession: A45063

R:Kihlberg, B.M.; Sjoerding, U.; Kastern, W.; Bjoerck, L.

J. Biol. Chem. 267, 25583-25588, 1992

A:Title: Protein LG: a hybrid molecule with unique immunoglobulin binding properties.

A:Reference number: A45063; MUID:93094283; PMID:1460053

A:Accession: A45063

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-455 <KIH>

A:Cross-references: GB:S50809; NID:g261705; PIDN:AAA03280.1; PID:g261706

A:Note: sequence extracted from NCBI backbone (NCBIN:120302, NCBIP:120303)

Query Match 100.0%; Score 389; DB 2; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.2e-31;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKONGEYT 60  
DB 26 KEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKONGEYT 85  
QY 61 DVADKGYTLNIKFPAG 76  
DB 86 DVADKGYTLNIKFPAG 101

### RESULT 2

A42808

Ig light chain-binding protein precursor - Peptostreptococcus magnus

N:Alternate names: protein L

C:Species: Peptostreptococcus magnus

C>Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 07-May-1999

C:Accession: A42808; A41493

R:Kastern, W.; Sjoerding, U.; Bjoerck, L.

J. Biol. Chem. 267, 12820-12825, 1992

A:Title: Structure of peptostreptococcal protein L and identification of a repeated i

A:Reference number: A42808; MUID:92316971; PMID:1618782

A:Accession: A42808

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-719 <KAS>

A:Cross-references: GB:M86697

R:Kastern, W.; Holst, E.; Nielsen, E.; Sjoerding, U.; Bjoerck, L.

Infect. Immun. 58, 1217-1222, 1990

A:Title: Protein L, a bacterial immunoglobulin-binding protein and possible virulence

A:Reference number: A41493; MUID:90215984; PMID:2108927

A:Accession: A41493

A:Status: preliminary





**A: Gene: NUKINI**

Argene: CESP:zk003.0  
 A:Map position: 5  
 A:Introns: 11/3

Query Match 17.78; Score 69; DB 2; Length 123;

Best Local Similarity 24.4%; Pred. No. 5.9;  
Matches 21; Conservative 10; Mismatches 31; Indels 24; Gaps 2;

QY 1 KEETPEPTP-----ETDSEEVTKANLIFANGSTQTAEFKGTPEKATSEAY 48  
| | | | | | | | | | : | | | | | : |  
Db 14 KMDTAEPAAAPRAASAAAPEAESNNTVTPSVNLGSGQQQTGNQSAPRNTSTVTROY 73  
| | | | | | | | | | : | | | | | : |

QY 49 ADT-----LKKNGEYTVTD 62  
| : | | | | | : |  
Db 74 LDSTVVPIILQLGLAKDRPENPIE 99  
| | | | | | | | | | : | | | | | : |

RESULT 12  
S67641  
hypothetical protein YDL099w - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein D2375  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
C:Accession: S67641; S67420  
R:Ballesta, J.P.G.; Remacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.; Bo submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67629  
A:Accession: S67641  
A:Molecule type: DNA  
A:Residues: 1-341 <BAL>  
A:Cross-references: EMBL:D74147; NID:g1431135; PIDN:CAA98666.1; PID:e253030; PID:g143133  
A:Experimental source: strain S288C  
R:Bozkovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jimenez submitted to the EMBL Data Library, February 1996  
A:Reference number: S67406  
A:Accession: S67420  
A:Molecule type: DNA  
A:Residues: 1-341 <BOS>  
A:Cross-references: EMBL:X95644; NID:g1199535; PIDN:CAA64914.1; PID:e223192; PID:g119955  
C:Genetics:  
A:Cross-references: SGD:S0002257  
A:Map position: 4L  
A>Note: YDL099w

Query Match 17.7%; Score 69; DB 2; Length 341;  
Best Local Similarity 25.0%; Pred. No. 18;  
Matches 25; Conservative 13; Mismatches 30; Indels 32; Gaps 3;

QY 3 ETPE---TPETDSEEVTKANLIFANGSTQ-----AFKGTTFEKATSE 44  
||| : | : | | | | : | : | : | : | : | : | : |  
Db 50 ETPDLTGPGESTQEETVKANSKSENNDVDDEESEKEIEQVKSDPSGTTEKDIE 109  
| | | | | | | | | | : | | | | | : |

QY 45 AYADATLKDNGEY-----VDVADKGYL 70  
: : | | | | : | : | : | : | : | : | : |  
Db 110 INSTSNVGKDDAENTKKEEVQVIKNNDNQADAGKTI 149  
| | | | | | | | | | : | | | | | : |

RESULT 13  
B82557  
hypothetical protein XF2445 [imported] - *Xylella fastidiosa* (strain 9a5c)  
C:Species: *Xylella fastidiosa*  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82557  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A:Reference number: AB3515; MUID:20355717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82557  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1056 <SIM>  
A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF85244.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canarog, L.E.A.; Carraro, D.M.; Carrier, H.; Neto, E.; Docena, C.; El-Dorri, H.; Pacinani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000

Search completed: September 3, 2003, 11:24:28  
Job time : 23.9589 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:52 ; Search time 9.63014 Seconds  
(without alignments)  
371.130 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEETPTETPDEEVTIKA.....GEYTDVADKGYTLNIKIFAG 76

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	73	18.8	634	1	GYRB_BORHE
2	69	17.7	123	1	DP30_CAEEL
3	68	17.5	574	1	TACY_STRCB
4	68	17.5	574	1	TACY_STRCB
5	67	17.2	571	1	TACY_STRPY
6	66.5	17.1	840	1	YHT1_YEAST
7	66	17.0	418	1	PGK_EUPCR
8	64.5	16.6	300	1	JAM1_MOUSE
9	64.5	16.6	446	1	GBAL_KLULA
10	64	16.5	544	1	CAR9_CANAL
11	63.5	16.3	442	1	INSG_ECOLI
12	63.5	16.3	1409	1	HAP1_HAEIN
13	63	16.2	815	1	C48D_ARATH
14	63	16.2	3562	1	PGCV_CHICK
15	62.5	16.1	190	1	HBPI_RHIAP
16	62.5	16.1	269	1	HTSL_STRFI
17	62.5	16.1	374	1	FCGI_HUMAN
18	62	15.9	199	1	PVAA_STRP8
19	62	15.9	644	1	YHOB_DROME
20	61.5	15.8	404	1	HUTI_TREVO
21	61.5	15.8	516	1	YDDS_ECOLI
22	61.5	15.8	646	1	DHSA_CAEEL
23	61.5	15.8	1139	1	LI25_CAEEL
24	61	15.7	311	1	RPOA_STRPN
25	61	15.7	688	1	EFG_CLOPE
26	61	15.7	1120	1	STFR_ECOLI
27	61	15.7	1167	1	SCA2_STRPY
28	61	15.7	1181	1	SCA2_STRPY
29	60.5	15.6	514	1	YJJI_HAEIN
30	60.5	15.6	613	1	DNAK_LACSK
31	60.5	15.6	625	1	HTPG_BUCAP
32	60.5	15.6	1176	1	SLAP_BACSH
33	60	15.4	497	1	FTSY_ECOLI

#### RESULT 1

GYRB\_BORHE 15.4 60 15.4 588 1 PPCK\_THEAC  
ID GYRB\_BORHE STANDARD; PRT; 634 AA.  
AC Q9ZFK1;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE DNA gyrase subunit B (EC 5.99.1.3).  
GN GYRB.  
OS Borrelia hermsii.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=140;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HS1;  
RA Samuels D.S., Kimmel B.J., Huang W.M.;  
RT "Mutations in Borrelia hermsii gyrb confer resistance to coumermycin  
Al.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE  
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED  
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
CC of double-stranded DNA.  
CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA  
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE  
CC ENZYME FORMS AN A2B2 TETRAMER.  
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF098862; AAC72846.1; -  
CC HSP; P06982; IAJ6.  
DR InterPro; IPR003594; ATPbind\_ATPase.  
DR InterPro; IPR002288; DNA\_gyraseB\_C.  
DR InterPro; IPR000565; DNA\_gyrb.  
DR InterPro; IPR001241; DNA\_topoisoII.  
DR InterPro; IPR006171; Toprim\_dom.  
DR Pfam; PF00204; DNA\_gyraseB; 1.  
DR Pfam; PF00986; DNA\_gyraseB\_C; 1.  
DR Pfam; PF02518; HATPase\_C; 1.  
DR Pfam; PF01751; Toprim; 1.  
DR PRINTS; PR00418; TP12FAMILY.  
DR PRODOM; PD149633; DNA\_gyraseB\_C; 1.  
DR SMART; SM00387; HATPase\_c; 1.  
DR SMART; SM00433; TOP2c; 1.  
DR TIGRFAMs; TIGR01059; gyrb; 1.  
DR PROSITE; PS00177; TOPOISOMERASE-II; 1.  
KW Topoisomerase; Isomerase; ATP-binding.

O9hlv2 thermoplas  
O9kka3 r outer mem  
P14873 mus musculu  
O9wu42 mus musculu  
P19158 saccharomyc  
Q8rc30 thermoanaer  
P08921 rattus norv  
P05423 homo sapien  
P35829 lactobacill  
P09913 homo sapien  
O43493 homo sapien  
O35608 mus musculu

23

```

GN Streptococcus equisimilis.
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SIMD-1;
RX MEDLINE=95102113; PubMed=7803818;
RA Okumura K., Hara A., Tanaka T., Nichiguchi I., Minamide W.,
RA Igarashi H., Yutsudo T.; the streptolysin O genes of group C and group
RT "Cloning and sequencing the streptolysin O genes of group C and group
RL G streptococci.";
RL DNA Seq. 4:325-328(1994).
CC -1- FUNCTION: SULPHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16824; BAA04104.1;
DR HSP; P19995; IPFO.
DR InterPro; IPR001869; Thiol_cytolysin.
DR Pfam; PF01289; Thiol_cytolysin; 1.
DR PRINTS; PR01400; TACYTOLYSIN.
DR ProDom; PD007062; Thiol_cytolysin; 1.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
KW Toxin; Hemolysis; Cytolysis; Signal.
FT SIGNAL 1 36 BY SIMILARITY.
FT CHAIN 37 574 STREPTOLYSIN O.
FT ACT_SITE 533 533 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 574 AA; 63992 MW; 83345CEBDE1EB9CFC CRC64;

Query Match 17.5%; Score 68; DB 1; Length 574;
Best Local Similarity 30.1%; Pred. No. 17;
Matches 22; Conservative 13; Mismatches 32; Indels 6; Gaps 2;

QY 4 TPETPTDSEEVTIKA-----NLIFANGSTQTAEEKGTPEKATSEAYAYADTLKKDNGE 58
DB 49 TNEQPKPESELTEKAGQKMDMLNSNDMIKAPKEMPLESAEKEKKSEDN-KKSEED 107
QY 59 YTVDVADKGYTLN 71
DB 108 HTEINDKIYSLN 120

RESULT 5
TACY_STRPY STANDARD; PRT; 571 AA.
AC P21131;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Streptolysin O precursor (Thiol-activated cytolysin).
GN SLO OR SPY0167 OR SPY0130 OR SPY0132.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88057628; PubMed=3502717;
RA Kehoe M.A., Miller L., Walker J.A., Boulnois G.J.;
RA "Nucleotide sequence of the streptolysin O (SLO) gene: structural

```

```

RT homologies between SLO and other membrane-damaging, thiol-activated
RT toxins."
RL Infect. Immun. 55:3228-3232(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Perretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SULPHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M18638; AAA26975.1;
DR EMBL; AE006485; AAK33267.1;
DR EMBL; AE014138; AAM78737.1;
DR EMBL; AP005141; BAC63227.1; ALT_INIT.
DR PIR; A43507; A43507.
DR HSP; P19995; IPFO.
DR InterPro; IPR001869; Thiol_cytolysin.
DR Pfam; PF01289; Thiol_cytolysin; 1.
DR PRINTS; PR01400; TACYTOLYSIN.
DR ProDom; PD007062; Thiol_cytolysin; 1.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
KW Toxin; Hemolysis; Cytolysis; Signal; Complete proteome.
FT SIGNAL 1 33
FT CHAIN 34 571 STREPTOLYSIN O.
FT ACT_SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 571 AA; 63638 MW; D05AA9979DCBA80 CRC64;

Query Match 17.2%; Score 67; DB 1; Length 571;
Best Local Similarity 30.1%; Pred. No. 22;
Matches 22; Conservative 13; Mismatches 32; Indels 6; Gaps 2;

QY 4 TPETPTDSEEVTIKA-----NLIFANGSTQTAEEKGTPEKATSEAYAYADTLKKDNGE 58
DB 46 TNEQPKPESELTEKAGQKTDMLNSNDMIKAPKEMPLESAEKEKKSEDK-KKSEED 104

```

[illegible]





```
CAR9_CANAL
ID CAR9_CANAL STANDARD; PRT; 544 AA.
AC Q42779;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Candidapepsin 9 precursor (EC 3.4.23.24) (Aspartate protease 9)
DE (ACP 9) (Secreted aspartic protease 9).
GN SAP9.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C74;
RX MEDLINE=99018822; PubMed=9802014;
RA Monod M., Hube B., Hess D., Sanglard D.;
RT "Differential regulation of SAP8 and SAP9, which encode two new
RT members of the secreted aspartic proteinase family in Candida
RT albicans.";
RL Microbiology 144:2731-2737(1998).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage at the carboxyl of
CC hydrophobic amino acids, but fails to cleave 15-Leu-|-Tyr-16, 16-
CC Tyr-|-Leu-17 and 24-Phe-|-Phe-25 of insulin B chain. Activates
CC trypsinogen, and degrades keratin.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF043331; AAC69996.1; -.
CC HSSP; P28871; IEAG.
CC MEROPS; A01.0PW; -.
DR InterPro; IPR001969; Aspprotease_site.
DR InterPro; IPR001461; AspproteaseA1.
DR Pfam; PF00026; asp; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 ? POTENTIAL.
FT CHAIN 2 544 CANDIDAPEPSIN 9.
FT ACT_SITE 167 167 BY SIMILARITY.
FT ACT_SITE 371 371 BY SIMILARITY.
FT DISULFID 406 441 BY SIMILARITY.
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 544 AA; 58624 MW; E6D755A46040A486 CRC64;

Query Match 16.5%; Score 64; DB 1; Length 544;
Best Local Similarity 27.6%; Pred. No. 41;
Matches 21; Conservative 18; Mismatches 19; Indels 18; Gaps 4;

QY 18 IKANLIFANGSTQTAFFKGT-----FEKATSEAYAYADTL-----KKDNGEY----- 60
Db 347 LKIDVSSSGST-TNLSGTTGWLDTGSLSVFSDTLQSLGKALNGQYSNVGAYVYN 405
QY 61 VDVAQGYTLNFKAG 76
Db 406 CNLADSSRTVDIEFG 421

CAR9_CANAL
ID INSG_ECOLI STANDARD; PRT; 442 AA.
AC P03835;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transposase insg for insertion sequence element IS4.
DE INSG OR B4278.
GN Escherichia coli.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=82012962; PubMed=6268937;
RA Klaer R., Kuhn S., Tillmann E., Fritz H.-J., Starlinger P.;
RT "The sequence of IS4.";
RL Mol. Gen. Genet. 181:169-175(1981).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
CC -|- FUNCTION: INVOLVED IN THE TRANSPOSITION OF THE INSERTION
CC SEQUENCE IS4.
CC -|- SIMILARITY: BELONGS TO THE TRANSPOSASE FAMILY 11.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; J01733; -. NOT_ANNOTATED_CDS.
CC EMBL; U14003; AAA97174.1; -.
CC EMBL; AE000498; AAC77234.1; -.
CC FIR; A04463; IEEC41.
CC EcoGene; EG40007; insG.
CC InterPro; IPR002559; Transposase_11.
CC Pfam; PF01609; Transposase_11; 1.
KW Transposase element; Transposition; DNA-binding; DNA recombination;
KW Complete proteome.
SQ SEQUENCE 442 AA; 50385 MW; 0B9CAAA0FFAC724E CRC64;

Query Match 16.3%; Score 63.5; DB 1; Length 442;
Best Local Similarity 25.6%; Pred. No. 37;
Matches 20; Conservative 14; Mismatches 27; Indels 17; Gaps 3;

QY 4 TPTETPDSE-----EVTIKANLIFANGSTQTAFFKGTFFKATSEAYAYAD 50
Db 139 TPTDTPDAAFPQTHAGNPALYPQVKVQCMELTSHLLTAAAF-GTMKNSENE---LAE 194

QY 51 TLKKNGEYTVDAQGY 68
Db 195 QLIEQTGDNLTLMDKGY 212

RESULT 12
ID HAP1_HAEIN STANDARD; PRT; 1409 AA.
AC P44596;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adhesion and penetration protein precursor (EC 3.4.21.-).
```



DR InterPro: IPR003338; ATPaseVAT\_N.  
 DR Pfam: PF00004; AAA; 2.  
 DR Pfam: PF02359; cdc48\_N; 1.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRPFAMs: TIGR01243; CDC48; 1.  
 DR PROSITE: PS00674; AAA; 2.  
 KW Hypothetical protein; Cell division; Cell cycle; Transport;  
 KW Protein transport; Nuclear protein; ATP-binding; Repeat;  
 KW Multigene family.  
 FT NP\_BIND 249 256 ATP (POTENTIAL).  
 FT NP\_BIND 522 529 ATP (POTENTIAL).  
 SQ SEQUENCE 815 AA; 90340 MW; 8F2817396748B6CC CRC64;  
 Query Match 16.2%; Score 63; DB 1; Length 815;  
 Best Local Similarity 28.6%; Pred. No. 80;  
 Matches 22; Conservative 12; Mismatches 27; Indels 16; Gaps 3;  
 QY 1 KEETPTPTDSEEVTKA-----NLIIFANGSTQAEFGKTFERATSEAYAYADTLKKD 55  
 Db 717 RAESPAMEDEETAEIAGHFESMKYARRSVSDAIR-----KYQAFQOTLQOS 768  
 QY 56 NG--EYTVDVADKGYT 69  
 Db 769 RGFSGSEFRPDAPTGT 785  
 RESULT 14  
 PGCV\_CHICK STANDARD; PRT; 3562 AA.  
 AC Q90953; Q90945;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Versican core protein precursor (Large fibroblast proteoglycan)  
 DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).  
 GN CP5G2.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxId=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).  
 RC STRAIN=White leghorn; TISSUE=Limb bud;  
 RX MEDLINE=93300846; PubMed=8314802;  
 RA Shinomura T., Nishida Y., Ito K., Kimata K.;  
 RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan  
 RT expressed during chondrogenesis in chick limb buds. Alternative  
 RT spliced multi-forms of PG-M and their relationships to versican.";  
 RL J. Biol. Chem. 268:14461-14469(1993).  
 CC -!- FUNCTION: May play a role in intercellular signaling and in  
 CC connecting cells with the extracellular matrix. May take part in  
 CC the regulation of cell motility, growth and differentiation. Binds  
 CC hyaluronic acid.  
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing: Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=V0;  
 CC IsoId=Q90953-1; Sequence=Displayed;  
 CC Name=V1;  
 CC IsoId=Q90953-2; Sequence=VSP\_003093;  
 CC -!- TISSUE SPECIFICITY: Prechondrogenic condensation area of  
 CC developing limb buds.  
 CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development  
 CC (By similarity).  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -!- SIMILARITY: Contains 2 link domains.  
 CC -!- SIMILARITY: Contains 2 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
 CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X60226; CAA42787.1; -.  
 CC EMBL; D13542; BAA02742.1; -.  
 CC PIR; A47171; A47171.  
 CC HSSP; P00740; 1EDM.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF\_like.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR InterPro: IPR000538; Link.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00047; ig; 1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sushi; 1.  
 DR Pfam: PF00193; Xlink; 2.  
 DR PRINTS; PR01265; LINKMODULE.  
 DR ProDom; PD000918; Link; 2.  
 DR SMART; SM00032; CCP; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00409; Ig; 1.  
 DR SMART; SM00445; LINK; 2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS01241; LINK; 2.  
 KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;  
 KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;  
 KW Hyaluronic acid; Alternative splicing.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 3562 VERSICAN CORE PROTEIN.  
 FT DOMAIN 27 143 IG-LIKE V-TYPE.  
 FT DOMAIN 166 243 LINK 1.  
 FT DOMAIN 264 345 LINK 2.  
 FT DOMAIN 3284 3290 EGF-LIKE 1.  
 FT DOMAIN 3292 3328 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 3341 3455 C-TYPE LECTIN.  
 FT DOMAIN 3460 3518 SUSHI.  
 FT DISULFID 44 129 BY SIMILARITY.  
 FT DISULFID 171 242 BY SIMILARITY.  
 FT DISULFID 195 216 BY SIMILARITY.  
 FT DISULFID 269 344 BY SIMILARITY.  
 FT DISULFID 293 314 BY SIMILARITY.  
 FT DISULFID 3258 3269 BY SIMILARITY.  
 FT DISULFID 3263 3278 BY SIMILARITY.  
 FT DISULFID 3280 3289 BY SIMILARITY.  
 FT DISULFID 3296 3307 BY SIMILARITY.  
 FT DISULFID 3301 3316 BY SIMILARITY.  
 FT DISULFID 3318 3327 BY SIMILARITY.  
 FT DISULFID 3334 3345 BY SIMILARITY.  
 FT DISULFID 3362 3454 BY SIMILARITY.  
 FT DISULFID 3430 3446 BY SIMILARITY.  
 FT DISULFID 3461 3504 BY SIMILARITY.  
 FT DISULFID 3490 3517 BY SIMILARITY.  
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 948 948 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1409 1409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1479 1479 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1530 1530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1625 1625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1988 1988 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2088 2088 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2089 2089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2507 2507 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2642 2642 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2679 2679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2748 2748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2762 2762 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3069 3069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3194 3194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3232 3232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3545 3545 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 485 1411 Missing (in isoform V1).
FT SQ SEQUENCE 3562 AA; 388078 MW; 9BC565E88C1602D2 CRC64;

Query Match 16.2%; Score 63; DB 1; Length 3562;
Best Local Similarity 28.6%; Pred. No. 3.9e+02;
Matches 26; Conservative 13; Mismatches 30; Indels 22; Gaps 5;

QY 6 ETPEPDSREVTIKANLIFANGSTQTAEFKGT--FEKAT-----SEAYAYADTLKKDNG 57
DB 863 ERAETESDEV-VSADFDTGKTGVTFVHTSSLDLEKFTLSKIPEDSSATVKFSFSSSG 921

QY 58 -----EYTVDVADK--GYTLNIRKA 75
DB 922 TVLPTAVATVLEVTDEHDETSGYVLNMTFS 952

RESULT 15
HBPI_RHIAP
ID HBPI_RHIAP STANDARD; PRT; 190 AA.
AC Q77420;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Female-specific histamine-binding protein 1 precursor (FS-HBPI).
OS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
OX NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=99288454; PubMed=10360182;
RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "Tick histamine-binding proteins: isolation, cloning, and three-
dimensional structure.";
RL Mol. Cell 3:661-671(1999).
CC -!- FUNCTION: BINDS HISTAMINE WITH A HIGH-AFFINITY. THE ABILITY TO
CC OUTCOMPETE HISTAMINE RECEPTORS INDICATES THAT ITS FUNCTION IS TO
CC SUPPRESS INFLAMMATION DURING BLOOD FEEDING.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U96080; AAC63106.1; -
DR HSSP: O77421; 1QFT.
DR InterPro: IPR002970; His_binding.
```

```
DR Pfam: PF02098; His_binding; 1.
DR ProDom: PD152455; His_binding; 1.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 190 FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN
FT 1.
FT DISULFID 66 187 BY SIMILARITY.
FT DISULFID 137 166 BY SIMILARITY.
SQ SEQUENCE 190 AA; 21370 MW; 855BE151A90053B1 CRC64;

Query Match 16.1%; Score 62.5; DB 1; Length 190;
Best Local Similarity 31.6%; Pred. No. 19;
Matches 18; Conservative 13; Mismatches 19; Indels 7; Gaps 3;

QY 12 SEEEVTIKANLIFANGSTQTAEFKGTFEKAT-SEAYAY----ADTLKKDNGEYTVDV 63
DB 75 NEDEKNVEAFWFMNNAIV--YQHTFEKATPKMYGKKNNAITYQTEDGQVLTDV 129

Search completed: September 3, 2003, 11:20:02
Job time : 11.6301 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:17 ; Search time 48.3288 Seconds  
(without alignments)  
236.470 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_81\_152

Perfect score: 371

Sequence: 1 KEKTEPEKPEVTKANLIY.....GEYTVDVADKGYTLNPKFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	371	100.0	72	21	AA1982538
2	371	100.0	305	14	AA1982538
3	371	100.0	367	21	AA1982538
4	371	100.0	434	14	AA1982538
5	371	100.0	467	22	AA1982538
6	352	94.9	72	21	AA1982538
7	322	86.8	72	21	AA1982538
8	314	84.6	76	21	AA1982538
9	308	83.0	75	21	AA1982538

10	308	83.0	291	14	AA1982538
11	308	83.0	1027	14	AA1982538
12	308	83.0	1027	14	AA1982538
13	299	80.6	182	18	AA1982538
14	299	80.6	182	18	AA1982538
15	292.5	78.8	71	21	AA1982538
16	291	78.4	178	18	AA1982538
17	291	78.4	178	18	AA1982538
18	291	78.4	198	18	AA1982538
19	291	78.4	198	18	AA1982538
20	291	78.4	342	18	AA1982538
21	291	78.4	342	18	AA1982538
22	291	78.4	482	20	AA1982538
23	291	78.4	495	18	AA1982538
24	291	78.4	495	18	AA1982538
25	289.5	78.0	71	21	AA1982538
26	286	77.1	82	21	AA1982538
27	283	76.3	74	21	AA1982538
28	279	75.2	82	21	AA1982538
29	279	75.2	82	21	AA1982538
30	277	74.7	82	21	AA1982538
31	273	73.6	82	21	AA1982538
32	90.5	24.4	395	22	ABG11860
33	89.5	24.1	463	22	ABG11860
34	73	19.7	588	22	ABG11860
35	69.5	18.7	279	23	ABG77964
36	68.5	18.5	261	20	AA1982538
37	68.5	18.5	344	17	AA1982538
38	68.5	18.5	374	17	AA1982538
39	68.5	18.5	374	19	AA1982538
40	68.5	18.5	374	19	AA1982538
41	68.5	18.5	374	19	AA1982538
42	68.5	18.5	374	21	AA1982538
43	68.5	18.5	399	21	AA1982538
44	68.5	18.5	399	22	AA1982538
45	67.5	18.2	344	13	AA1982538

#### ALIGNMENTS

RESULT 1  
AA1982538  
ID AA1982538 standard; Protein; 72 AA.

XX  
AC AA1982538;

DT  
XX 20-JUL-2000 (first entry)

XX  
DE Peptostreptococcus strain 312 protein L domain B2 protein sequence.

XX  
KW Immunoglobulin light chain binding protein; PpL; protein L;  
KW Peptostreptococcus; human immunoglobulin kappa chain;  
KW immunoaffinity chromatography.

XX  
OS Peptostreptococcus sp.

XX  
PN WO200015803-A1.

XX  
PD 23-MAR-2000.

XX  
PF 14-SEP-1999; 99WO-GB03048.

XX  
PR 14-SEP-1998; 98GB-0019998.

XX  
PR 26-APR-1999; 99GB-0009578.

XX  
(ACTI-) ACTINOVA LTD.

PI  
Gore MG, Beckingham JA, Roberts SE;

XX  
WPI; 2000-271441/23.

DR  
N-PSDB; AAA08427.

XX





44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873

PA . (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM..

XX Shani Z, Shoseyov O;  
 XX WPI; 2001-112219/12.  
 DR N-PSDB; AAF24730.  
 XX  
 XX Expressing and isolating recombinant protein in a plant, useful for  
 PT producing large quantities of recombinant proteins, by expressing a  
 PT fusion protein including a cellulose binding peptide fused to a  
 PT recombinant protein -  
 XX  
 XX Example; Fig 2a; 87pp; English.  
 XX  
 CC The specification describes a method for expressing and isolating  
 CC a recombinant protein in a plant. The method comprising expressing a  
 CC fusion protein including the recombinant protein and a cellulose  
 CC binding peptide fused to it, where the fusion protein is  
 CC compartmentalised and sequestered within plant cells, plant derived  
 CC tissue or cultured plant cells. The method is useful for obtaining large  
 CC quantities of the recombinant proteins and protein products in a simple  
 CC and cost-effective manner. Recombinant proteins may be used commercially,  
 CC such as in the food processing industry, e.g. glucosylases and glucose  
 CC isomerases are used for converting starch to high fructose corn syrup,  
 CC proteinases for the hydrolysis of high molecular weight proteins and in  
 CC manufacturing leather or alcoholic beverages, pectinesterases for  
 CC pectin hydrolysis in food industry, lipases for cleaving ester linkage  
 CC in triglycerides, and for effluent treatment. The recombinant proteins  
 CC may further be used to produce protein antibiotics, which can be used  
 CC in healing processes, and to produce animal feed enzymes. The present  
 CC sequence represents a fusion protein of the invention, and comprises a  
 CC fusion of a cell signal peptide, protein L, CBD cex, and an endoplasmic  
 CC reticulum retaining peptide.  
 XX  
 SQ Sequence 467 AA;  
 Query Match 100.0%; Score 371; DB 22; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-36;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KETPEPEKEEVTIKANLIYADGKTQTAEFGTFFETAEAYRYADALKKDNGEYTVDA 60  
 Db 120 KETPEPEKEEVTIKANLIYADGKTQTAEFGTFFETAEAYRYADALKKDNGEYTVDA 179  
 QY 61 DKGYTLNIRFAG 72  
 Db 180 DKGYTLNIRFAG 191  
 RESULT 6  
 AAY82539  
 ID AAY82539 standard; Protein; 72 AA.  
 AC AAY82539;  
 XX  
 XX 20-JUL-2000 (first entry)  
 DT  
 DE Peptostreptococcus strain 312 protein L domain B3 protein sequence.  
 XX  
 KW Immunoglobulin light chain binding protein; PpL; protein L;  
 KW Peptostreptococcus; human immunoglobulin kappa chain;  
 KW Immunoaffinity chromatography.  
 XX  
 OS Peptostreptococcus sp.  
 XX  
 PN WO200015803-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 14-SEP-1999; 99WO-GB03048.  
 XX  
 PR Immunoglobulin light chain binding protein; PpL; protein L;  
 PR Peptostreptococcus; human immunoglobulin kappa chain;  
 PR Immunoaffinity chromatography.  
 XX  
 OS Peptostreptococcus sp.  
 XX  
 PN WO200015803-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 14-SEP-1999; 99WO-GB03048.  
 XX  
 PR 14-SEP-1998; 98GB-0019998.  
 PR 26-APR-1999; 99GB-0009578.  
 XX  
 XX

PA (ACTI-) ACTINOVA LTD.  
 XX  
 XX Gore MG, Beckingham JA, Roberts SE;  
 DR WPI; 2000-271441/23.  
 DR N-PSDB; AAA08428.  
 XX  
 XX New modified immunoglobulin light chain binding protein, useful in  
 PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
 PT more at pH8 with respect to human immunoglobulin kappa-chain -  
 XX  
 XX Disclosure; Page 43; 56pp; English.  
 XX  
 CC The present invention describes an immunoglobulin (Ig) light chain  
 CC binding protein (P1) which has been modified by one or more amino acid  
 CC substitutions such that the dissociation constant (Kd) of the protein  
 CC with respect to human Ig kappa-chain is 400 nM or more at pH8. P1 is  
 CC useful in immunoaffinity chromatography. The present sequence is a  
 CC Peptostreptococcus protein L Ig light chain binding domain, which is  
 CC given in the disclosure of the present invention.  
 XX  
 SQ Sequence 72 AA;  
 Query Match 94.9%; Score 352; DB 21; Length 72;  
 Best Local Similarity 94.4%; Pred. No. 4.8e-35;  
 Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KETPEPEKEEVTIKANLIYADGKTQTAEFGTFFETAEAYRYADALKKDNGEYTVDA 60  
 Db 1 KETPEPEKEEVTIKANLIYADGKTQTAEFGTFFETAEAYRYADALKKDNGEYTVDA 60  
 QY 61 DKGYTLNIRFAG 72  
 Db 61 DKGYTLNIRFAG 72  
 RESULT 7  
 AAY82540  
 ID AAY82540 standard; Protein; 72 AA.  
 AC AAY82540;  
 XX  
 XX 20-JUL-2000 (first entry)  
 DT  
 DE Peptostreptococcus strain 312 protein L domain B4 protein sequence.  
 XX  
 KW Immunoglobulin light chain binding protein; PpL; protein L;  
 KW Peptostreptococcus; human immunoglobulin kappa chain;  
 KW Immunoaffinity chromatography.  
 XX  
 OS Peptostreptococcus sp.  
 XX  
 PN WO200015803-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 14-SEP-1999; 99WO-GB03048.  
 XX  
 PR 14-SEP-1998; 98GB-0019998.  
 PR 26-APR-1999; 99GB-0009578.  
 XX  
 XX (ACTI-) ACTINOVA LTD.  
 XX  
 XX Gore MG, Beckingham JA, Roberts SE;  
 DR WPI; 2000-271441/23.  
 DR N-PSDB; AAA08429.  
 XX  
 XX New modified immunoglobulin light chain binding protein, useful in  
 PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
 PT more at pH8 with respect to human immunoglobulin kappa-chain -  
 XX  
 XX Disclosure; Page 44-45; 56pp; English.  
 XX

XX The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (PI) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
SQ Sequence 72 AA;

Query Match 86.8%; Score 322; DB 21; Length 72;  
Best Local Similarity 84.7%; Pred. No. 2.1e-31;  
Matches 61; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADALKKNGEYTVDA 60  
DB 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADALKKNGEYTVDA 60  
QY 61 DRGYTLNIRFAG 72  
DB 61 DGYTLNIRFAG 72

RESULT 8  
AAY82537  
ID AAY82537 standard; Protein; 76 AA.  
XX  
AC AAY82537;  
XX  
DT 20-JUL-2000 (first entry)  
XX  
DE Peptostreptococcus strain 312 protein L domain B1 protein sequence.  
XX  
KW Immunoglobulin light chain binding protein; PpL; protein L;  
KW Peptostreptococcus; human immunoglobulin kappa chain;  
KW immunoaffinity chromatography.  
XX  
OS Peptostreptococcus sp.  
XX  
PN WO200015803-A1.  
XX  
PD 23-MAR-2000.  
XX  
PF 14-SEP-1999; 99WO-GB03048.  
XX  
PR 14-SEP-1998; 98GB-0019998.  
PR 26-APR-1999; 99GB-0009578.  
XX  
PA (ACTI-) ACTINOVA LTD.  
XX  
PI Gore MG, Beckingham JA, Roberts SE;  
XX  
DR WPI; 2000-271441/23.  
DR N-PSDB; AAA08426.  
XX  
PT New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
PT more at pH8 with respect to human immunoglobulin kappa-chain -  
XX  
PS Disclosure; Page 40-41; 56pp; English.  
XX  
CC The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (PI) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
SQ Sequence 76 AA;

Query Match 84.6%; Score 314; DB 21; Length 76;

Best Local Similarity 81.6%; Pred. No. 2.1e-30;  
Matches 62; Conservative 6; Mismatches 4; Indels 4; Gaps 1;

QY 1 KKTPEEP---KEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADALKKNGEYTV 56  
DB 1 KKTPEETPTDSEEEVTIKANLIFANGSTQTAEFGKTFEATSEAYAYADTLKKNGEYTV 60  
QY 57 VDVAADGKYTLNIRFAG 72  
DB 61 VDVAADGKYTLNIRFAG 76

RESULT 9  
AAY82544  
ID AAY82544 standard; Protein; 75 AA.  
XX  
AC AAY82544;  
XX  
DT 20-JUL-2000 (first entry)  
XX  
DE Peptostreptococcus strain 3316 protein L domain C4 protein sequence.  
XX  
KW Immunoglobulin light chain binding protein; PpL; protein L;  
KW Peptostreptococcus; human immunoglobulin kappa chain;  
KW immunoaffinity chromatography.  
XX  
OS Peptostreptococcus sp.  
XX  
PN WO200015803-A1.  
XX  
PD 23-MAR-2000.  
XX  
PF 14-SEP-1999; 99WO-GB03048.  
XX  
PR 14-SEP-1998; 98GB-0019998.  
PR 26-APR-1999; 99GB-0009578.  
XX  
PA (ACTI-) ACTINOVA LTD.  
XX  
PI Gore MG, Beckingham JA, Roberts SE;  
XX  
DR WPI; 2000-271441/23.  
DR N-PSDB; AAA08433.  
XX  
PT New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
PT more at pH8 with respect to human immunoglobulin kappa-chain -  
XX  
PS Disclosure; Page 49-50; 56pp; English.  
XX  
CC The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (PI) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
SQ Sequence 75 AA;

Query Match 83.0%; Score 308; DB 21; Length 75;  
Best Local Similarity 84.3%; Pred. No. 1.1e-29;  
Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 KPTPEEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADALKKNGEYTVDAVK 62  
DB 5 ETPEEPKEEVTIKANLIFADGKTQTAEFGKTFEATAEAYRYADLLAKVNGEYTVADLG 64  
QY 63 GYTLNIRFAG 72  
DB 65 GYTLNIRFAG 74

```
RESULT 10
AAR42204
ID AAR42204 standard; Protein; 291 AA.
XX AC
XX AAR42204;
XX AC
XX 25-MAR-2003 (updated)
XX 18-MAY-1994 (first entry)
XX DE
XX Immunoglobulin binding protein derived from protein L.
XX KW
XX Peptide; immunoglobulin; binding; analysis; purification; ELISA;
XX enzyme linked immunoabsorbant assay.
XX OS
XX Synthetic.
XX PN
XX WO9322439-A1.
XX 11-NOV-1993.
XX PF
XX 07-MAY-1993; 93WO-GB00950.
XX PR
XX 07-MAY-1992; 92GB-0009804.
XX 24-DEC-1992; 92GB-0026928.
XX PA
XX (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX PI
XX Atkinson A, Duggieby CJ, Murphy JP, Trowern AR;
XX WPI; 1993-368798/46.
XX DR
XX N-PSDB; AAQ50947.
XX PS
XX New immunoglobulin binding proteins derived from Protein L -
XX which bind immunoglobulin kappa light chains but not albumin or
XX cell walls
XX Claim 12; Figure 2; 28pp; English.
XX CC
XX The synthetic immunoglobulin binding proteins derived from protein
XX L comprise repeated sequences from protein L which bind
XX immunoglobulin kappa light chains. They can be used in protein
XX analysis, purification procedures and other biochemical processes e.
XX g. ELISA. The synthetic molecules are of particular advantage if
XX they are free of regions in protein L which exhibit albumin and cell
XX wall binding.
XX CC
XX (Updated on 25-MAR-2003 to correct PN field.)
XX SQ
XX Sequence 291 AA;
XX Query Match 83.0%; Score 308; DB 14; Length 291;
XX Best Local Similarity 84.3%; Pred. No. 6.5e-29;
XX Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
XX QY 3 KTEPEPKEVTIKANLIYADGKTQTAEFGTTEATAEAYRYADALKDNGEYTVDVADK 62
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 221 ETPEEPKEVTIKVNLIFADGKTQTAEFGTTEATAEAYRYADLLAKVNGEYTDLEGG 280.
XX Db
XX 63 GYTLNIKFPAG 72
XX |||:|||||
XX 281 GYTLNIKFPAG 290
XX Db
XX RESULT 11
XX AAR42203
XX ID AAR42203 standard; Protein; 1027 AA.
XX AC
XX AAR42203;
XX XX
XX 25-MAR-2003 (updated)
XX 18-MAY-1994 (first entry)
XX DT
XX Protein L.
XX DE
XX Peptide; immunoglobulin; binding; immobilisation; light chains;
XX antibodies; diagnosis; pharmaceutical; ss.
XX OS
XX Peptococcus magnus.
XX XX

KW Peptide; immunoglobulin; binding; analysis; purification; ELISA;
KW enzyme linked immunoabsorbant assay.
XX OS
XX Peptococcus magnus.
XX FH
XX Location/Qualifiers
XX FT
XX 36..59
XX /label= Signal sequence.
XX FT
XX 60..968
XX /label= Mature protein L.
XX PN
XX WO9322439-A1.
XX XX
XX 11-NOV-1993.
XX XX
XX 07-MAY-1993; 93WO-GB00950.
XX PR
XX 07-MAY-1992; 92GB-0009804.
XX 24-DEC-1992; 92GB-0026928.
XX XX
XX (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX PA
XX Atkinson A, Duggieby CJ, Murphy JP, Trowern AR;
XX PI
XX WPI; 1993-368798/46.
XX DR
XX N-PSDB; AAQ50946.
XX XX
XX New immunoglobulin binding proteins derived from Protein L -
XX which bind immunoglobulin kappa light chains but not albumin or
XX cell walls
XX PS
XX Disclosure; Figure 1; 28pp; English.
XX CC
XX The synthetic immunoglobulin binding proteins derived from protein
XX L correspond to the repeated sequences in protein L which bind
XX immunoglobulin kappa light chains. They can be used in protein
XX analysis, purification procedures and other biochemical processes e.
XX g. ELISA. The synthetic molecules are of particular advantage if
XX they are free of regions in protein L which exhibit albumin and cell
XX wall binding.
XX CC
XX (Updated on 25-MAR-2003 to correct PN field.)
XX XX
XX Sequence 1027 AA;
XX Query Match 83.0%; Score 308; DB 14; Length 1027;
XX Best Local Similarity 84.3%; Pred. No. 3.5e-28;
XX Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
XX QY 3 KTEPEPKEVTIKANLIYADGKTQTAEFGTTEATAEAYRYADALKDNGEYTVDVADK 62
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 503 ETPEEPKEVTIKVNLIFADGKTQTAEFGTTEATAEAYRYADLLAKVNGEYTDLEGG 562
XX Db
XX 63 GYTLNIKFPAG 72
XX |||:|||||
XX 563 GYTLNIKFPAG 572
XX Db
XX RESULT 12
XX AAR43699
XX ID AAR43699 standard; Protein; 1027 AA.
XX AC
XX AAR43699;
XX XX
XX 25-MAR-2003 (updated)
XX 18-MAY-1994 (first entry)
XX DT
XX Protein L.
XX DE
XX Peptide; immunoglobulin; binding; immobilisation; light chains;
XX antibodies; diagnosis; pharmaceutical; ss.
XX KW
XX Peptococcus magnus.
XX OS
XX Peptococcus magnus.
XX XX
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:16:02 ; Search time 15.7808 Seconds  
(without alignments)  
193.043 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_81\_152

Perfect score: 371

Sequence: 1 KEKTPPEKEVTIKANLIY.....GEYTVADVADKGYTLNIKTAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_AA:\*
- 1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep:\*
  - 2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep:\*
  - 3: /cgn2.6/ptodata/1/1aa/6A\_COMB.pep:\*
  - 4: /cgn2.6/ptodata/1/1aa/6B\_COMB.pep:\*
  - 5: /cgn2.6/ptodata/1/1aa/PCTUS\_COMB.pep:\*
  - 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371	100.0	305	2 US-08-795-475-1	Sequence 1, Appli
2	371	100.0	434	2 US-08-795-475-3	Sequence 3, Appli
3	308	83.0	75	3 US-08-446-137B-8	Sequence 8, Appli
4	308	83.0	291	3 US-08-446-137B-4	Sequence 4, Appli
5	308	83.0	1027	3 US-08-446-137B-2	Sequence 2, Appli
6	299	80.6	182	3 US-08-828-741B-2	Sequence 2, Appli
7	299	80.6	182	4 US-09-160-567-2	Sequence 2, Appli
8	299	80.6	182	4 US-09-710-299-2	Sequence 2, Appli
9	292.5	78.8	71	3 US-08-446-137B-6	Sequence 6, Appli
10	291	78.4	178	3 US-08-828-741B-13	Sequence 13, Appli
11	291	78.4	178	4 US-09-160-567-13	Sequence 13, Appli
12	291	78.4	178	4 US-09-710-299-13	Sequence 13, Appli
13	291	78.4	198	3 US-08-828-741B-8	Sequence 8, Appli
14	291	78.4	198	4 US-09-160-567-8	Sequence 8, Appli
15	291	78.4	198	4 US-09-710-299-8	Sequence 8, Appli
16	291	78.4	342	3 US-08-828-741B-6	Sequence 6, Appli
17	291	78.4	342	4 US-09-160-567-6	Sequence 6, Appli
18	291	78.4	342	4 US-09-710-299-6	Sequence 6, Appli
19	291	78.4	495	3 US-08-828-741B-4	Sequence 4, Appli
20	291	78.4	495	4 US-09-160-567-4	Sequence 4, Appli
21	291	78.4	495	4 US-09-710-299-4	Sequence 4, Appli
22	288.5	77.8	71	3 US-08-446-137B-5	Sequence 5, Appli
23	283	76.3	74	3 US-08-446-137B-7	Sequence 7, Appli
24	69.5	18.7	279	4 US-09-314-701-60	Sequence 60, Appli
25	66.5	17.9	138	6 5218099-1	Patent No. 5218099
26	65	17.5	419	4 US-09-134-001C-3441	Sequence 3441, Ap
27	64	17.3	262	3 US-08-961-0083-98	Sequence 98, Appli

ALIGNMENTS

RESULT 1

US-08-795-475-1

; Sequence 1, Application US/08795475

; Patent No. 5965390

; GENERAL INFORMATION:

; APPLICANT: Bjvick, Lars

; APPLICANT: Sjvbring, Ulf

; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED AND BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/795,475

; FILING DATE: 11-FEB-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Mcmasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 100084.402D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 305 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

US-08-795-475-1

Query Match 100.0%; Score 371; DB 2; Length 305;

Best Local Similarity 100.0%; Pred. No. 3.4e-38;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 KEKTPPEKEVTIKANLIYADGKTQTAEFGTFEATAAYADALKKNGEYTVDA 60

|||||

Db 81 KKTPEEPKEVTIRANLIYADGKTQTAEFKGTFEATAEAYRYADALKKONGEYTVDA 140

QY 61 DRGYTLINIKFAG 72

Db 141 DRGYTLINIKFAG 152

# RESULT 2

US-08-795-475-3

; Sequence 3, Application US/08/795475

; Patent No. 5965390

; GENERAL INFORMATION:

; APPLICANT: Bjvrck, Lars

; APPLICANT: Sjvbring, Ulf

; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/795,475

; FILING DATE: 11-FEB-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Mcmasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 100084.402D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 434 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055

US-08-795-475-3

Query Match 100.0%; Score 371; DB 2; Length 434;

Best Local Similarity 100.0%; Pred. No. 5.5e-38;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEVTIRANLIYADGKTQTAEFKGTFEATAEAYRYADALKKONGEYTVDA 60

Db 81 KKTPEEPKEVTIRANLIYADGKTQTAEFKGTFEATAEAYRYADALKKONGEYTVDA 140

QY 61 DRGYTLINIKFAG 72

Db 141 DRGYTLINIKFAG 152

# RESULT 3

US-08-446-137B-8

; Sequence 8, Application US/08446137B

; Patent No. 6162903

; GENERAL INFORMATION:

; APPLICANT: Trowen, Angus R.

; APPLICANT: Atkinson, Anthony

; APPLICANT: Murphy, Jonathan P.

; APPLICANT: Laurence, Oliver S.

; APPLICANT: Duggleby, Clive J.

; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED

; FROM L PROTEIN AND THEIR USES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,137B

; FILING DATE: 22-MAY-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Mcmasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 100084.406

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 75 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-446-137B-8

Query Match 83.0%; Score 308; DB 3; Length 75;

Best Local Similarity 84.3%; Pred. No. 3.6e-31;

Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 KTTPEEPKEVTIRANLIYADGKTQTAEFKGTFEATAEAYRYADALKKONGEYTVDA 62

Db 5 ETPEEPKEVTIKVNLIFADGKTQTAEFKGTFEATAEAYRYADLLAKVNGEYTDLEDG 64

QY 63 GYTNIKFAG 72

Db 65 GYTNIKFAG 74

# RESULT 4

US-08-446-137B-4

; Sequence 4, Application US/08446137B

; Patent No. 6162903

; GENERAL INFORMATION:

; APPLICANT: Trowen, Angus R.

; APPLICANT: Atkinson, Anthony

; APPLICANT: Murphy, Jonathan P.

; APPLICANT: Laurence, Oliver S.

; APPLICANT: Duggleby, Clive J.

; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED

; FROM L PROTEIN AND THEIR USES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30



```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMASTERS, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-137B-4

Query Match      83.0%; Score 308; DB 3; Length 291;
Best Local Similarity 84.3%; Pred. No. 2.2e-30;
Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 KTPPEPKEEVTIKANLIYADGKTQTAEFKGTFFETAEAYRYADALKKNGEYTVDVADK 62
Db 221 ETPEEPKEEVTIKVNLIFADGKTQTAEFKGTFFETAEAYRYADLLAKVNGEYTVADLE 280
QY 63 GYTINIKFAG 72
Db 281 GYTINIKFAG 290

RESULT 5
US-08-446-137B-2
; Sequence 2, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trower, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Dugleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMASTERS, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B-2
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMASTERS, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-137B-2

Query Match      83.0%; Score 308; DB 3; Length 291;
Best Local Similarity 84.3%; Pred. No. 2.2e-30;
Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 KTPPEPKEEVTIKANLIYADGKTQTAEFKGTFFETAEAYRYADALKKNGEYTVDVADK 62
Db 221 ETPEEPKEEVTIKVNLIFADGKTQTAEFKGTFFETAEAYRYADLLAKVNGEYTVADLE 280
QY 63 GYTINIKFAG 72
Db 281 GYTINIKFAG 290

RESULT 6
US-08-828-741B-2
; Sequence 2, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-828-741B-2

Query Match      80.6%; Score 299; DB 3; Length 182;
Best Local Similarity 80.3%; Pred. No. 1.5e-29;
Matches 57; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 EKTPEEPKEEVTIKANLIYADGKTQTAEFKGTFFETAEAYRYADALKKNGEYTVDVAD 61
Db 20 QAAPKONTEEVTIKANLIYADGKTQTAEFKGTFFETAEAYRYADLLAKVNGEYTVDVAD 79
QY 62 KGYTLNKFAG 72
Db 80 KGYTLNKFAG 90
```

RESULT 8  
US-09-710-299-2  
; Sequence 2, Application US/09710299  
; Patent No. 6521741  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; Suess, Gabriele M.  
; Tarlington, David M.  
; Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:

RESULT 9  
US-08-446-137B-6  
; Sequence 6, Application US/08446137B  
; Patent No. 6162903  
; GENERAL INFORMATION:  
; APPLICANT: Townert, Angus R.  
; APPLICANT: Atkinson, Anthony  
; APPLICANT: Murphy, Jonathan P.  
; APPLICANT: Laurence, Oliver S.  
; APPLICANT: Dugleby, Clive J.  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; FROM L PROTEIN AND THEIR USES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; MOLECULE TYPE: protein
; US-08-828-741b-13

Query Match      78.4%; Score 291; DB 3; Length 178;
Best Local Similarity 88.9%; Pred. No. 1.5e-28;
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKGTOTAEFKGTFEEATAEAYRYADALKKDNGEYTVDVADKGYTLNIK 69
Db 12 EEVTKANLIYADGKGTOTAEFKGTFEEATAEAYRYADALKKDNGEYTVDVADKGYTLNIK 71
QY 70 FAG 72
Db 72 FAG 74

RESULT 11
US-09-160-567-13
; Sequence 13, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-160-567-13

Query Match      78.4%; Score 291; DB 4; Length 178;
Best Local Similarity 88.9%; Pred. No. 1.5e-28;
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKGTOTAEFKGTFEEATAEAYRYADALKKDNGEYTVDVADKGYTLNIK 69
Db 12 EEVTKANLIYADGKGTOTAEFKGTFEEATAEAYRYADALKKDNGEYTVDVADKGYTLNIK 71
QY 70 FAG 72
Db 72 FAG 74

```

Db 72 FAG 74

## RESULT 12

US-09-710-299-13  
; Sequence 13, Application US/09710299

; Patent No. 6521741

## ; GENERAL INFORMATION:

; APPLICANT: Koentgen, Frank  
; Suess, Gabriele M.  
; Tarlington, David M.  
; Treutlein, Herbert R.

; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF

; PRODUCING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY &amp; PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09710,299

; FILING DATE: 09-No. 6521741-2000

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/828,741

; FILING DATE: &lt;Unknown&gt;

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 10591

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 178 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-710-299-13

Query Match

; Sequence 8, Application US/0828741B

; Best Local Similarity 88.9%; Pred. No. 1.5e-28;

Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKTQTAEFKGTPEEATAEAYRYADALKKNGEYTVADVADKGYTLNIK 69

Db 12 EEVTKANLIYADGKTQTAEFKGTPEEATAEAYRYADALKKNGEYTVADVADKGYTLNIK 71

QY 70 FAG 72

Db 72 FAG 74

## RESULT 13

US-08-828-741B-8

; Sequence 8, Application US/0828741B

; Patent No. 6043069

## ; GENERAL INFORMATION:

; APPLICANT: Koentgen, Frank

; Suess, Gabriele M.

; Tarlington, David M.

; Treutlein, Herbert R.

; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF

; PRODUCING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY &amp; PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: 08/828,741

; FILING DATE: 09-No. 6521741-2000

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/828,741

; FILING DATE: &lt;Unknown&gt;

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 10591

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 178 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-710-299-13

Query Match

; Sequence 8, Application US/09160567

; Patent No. 6326179

; GENERAL INFORMATION:

; APPLICANT: Koentgen, Frank

; Suess, Gabriele M.

; Tarlington, David M.

; Treutlein, Herbert R.

; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF

; PRODUCING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY &amp; PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: 08/828,741B

; FILING DATE: 26-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 10591

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 198 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-828-741B-8

Query Match

; Sequence 291; DB 3; Length 198;

; Best Local Similarity 88.9%; Pred. No. 1.7e-28;

Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKTQTAEFKGTPEEATAEAYRYADALKKNGEYTVADVADKGYTLNIK 69

Db 32 EEVTKANLIYADGKTQTAEFKGTPEEATAEAYRYADALKKNGEYTVADVADKGYTLNIK 91

QY 70 FAG 72

Db 92 FAG 94

RESULT 14

US-09-160-567-8

; Sequence 8, Application US/09160567

; Patent No. 6326179

; GENERAL INFORMATION:

; APPLICANT: Koentgen, Frank

; Suess, Gabriele M.

; Tarlington, David M.

; Treutlein, Herbert R.

; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF

; PRODUCING SAME

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY &amp; PRESSER

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: 08/828,741B

; FILING DATE: 26-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 10591

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 198 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-828-741B-8

APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-8

Query Match 78.4%; Score 291; DB 4; Length 198;  
Best Local Similarity 88.9%; Pred. No. 1.7e-28;  
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 10 EEVTKANLIYADGKTQTAEFKGTFFEEATAEAYRYADALKKDNGETVDVADKGYTLNIK 69  
DB 32 EEVTKANLIFANGSTQTAEFKGTFFEKATSEAYAYADTLKKNGETVDVADKGYTLNIK 91  
QY 70 FAG 72  
DB 92 FAG 94

## RESULT 15

US-09-710-299-8  
Sequence 8, Application US/09710299  
Patent No. 6521741  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
Suess, Gabriele M.  
Tarlington, David M.  
Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/710,299  
FILING DATE: 09-NO. 6521741-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-710-299-8  
Query Match 78.4%; Score 291; DB 4; Length 198;  
Best Local Similarity 88.9%; Pred. No. 1.7e-28;  
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 10 EEVTKANLIYADGKTQTAEFKGTFFEEATAEAYRYADALKKDNGETVDVADKGYTLNIK 69  
DB 32 EEVTKANLIFANGSTQTAEFKGTFFEKATSEAYAYADTLKKNGETVDVADKGYTLNIK 91  
QY 70 FAG 72  
DB 92 FAG 94

Search completed: September 3, 2003, 11:25:36  
Job time : 16.7808 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:23:22 ; Search time 26.6301 Seconds  
(without alignments)  
371.228 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_81\_152

Perfect score: 371  
Sequence: 1 KKTPEPKKEVTIKANLIY.....GEYTVDAKGYTLNKFAG 72

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications\_AA.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371	100.0	72	10	US-09-808-212A-6
2	371	100.0	305	8	US-08-325-278-1
3	371	100.0	434	8	US-08-325-278-3
4	352	94.9	72	10	US-09-808-212A-8
5	322	86.8	72	10	US-09-808-212A-10
6	314	84.6	76	10	US-09-808-212A-4
7	308	83.0	75	10	US-09-808-212A-18
8	299	80.6	182	12	US-10-345-618-2
9	292.5	78.8	71	10	US-09-808-212A-14
10	291	78.4	178	12	US-10-345-618-13
11	291	78.4	198	12	US-10-345-618-8
12	291	78.4	342	12	US-10-345-618-6
13	291	78.4	482	12	US-10-345-618-16
14	291	78.4	495	12	US-10-345-618-4
15	289.5	78.0	71	10	US-09-808-212A-12

16	286	77.1	82	10	US-09-808-212A-2
17	283	76.3	74	10	US-09-808-212A-16
18	69.5	18.7	279	14	US-10-059-964-60
19	69.5	18.7	279	15	US-10-314-639-60
20	68.5	18.5	261	10	US-09-245-764-7
21	68.5	18.5	399	9	US-09-925-301-1128
22	68.5	18.5	399	15	US-10-106-698-5782
23	65.5	17.7	829	9	US-09-815-242-11216
24	64	17.3	262	9	US-09-765-272-98
25	64	17.3	270	9	US-09-765-272-206
26	64	17.3	448	12	US-10-342-224-82
27	63.5	17.1	1588	12	US-10-154-419-37
28	63.5	17.1	1588	14	US-10-002-769-5
29	63	17.0	179	10	US-09-738-626-4795
30	62	16.7	841	9	US-09-815-242-5779
31	62	16.7	841	9	US-09-815-242-12751
32	61	16.4	836	10	US-09-858-525A-10
33	61	16.4	871	10	US-09-858-525A-2
34	60.5	16.3	109	10	US-09-816-989A-7
35	60.5	16.3	271	15	US-10-138-505-30
36	60.5	16.3	274	15	US-10-138-505-26
37	60.5	16.3	619	11	US-09-882-774-1
38	60.5	16.3	637	9	US-09-815-242-12058
39	60.5	16.3	637	15	US-10-127-032-168
40	60	16.2	1001	15	US-10-128-714-3240
41	59.5	16.0	455	9	US-09-191-989-4
42	59	15.9	102	15	US-10-016-516-12
43	59	15.9	300	10	US-09-953-499-10
44	59	15.9	1098	10	US-09-797-862-32
45	58.5	15.8	191	9	US-09-815-242-14038

ALIGNMENTS

RESULT 1  
US-09-808-212A-6  
; Sequence 6, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.41AUS  
; CURRENT APPLICATION NUMBER: US/09/808.212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-6

Query Match	100.0%	Score 371;	DB 10;	Length 72;
Best Local Similarity	100.0%	Pred. No. 1.6e-35;		
Matches 72;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	KKTPEPKKEVTIKANLIYADGKTQTAEFKGTFEETATAEAYADALKKNGEYTVDA 60		
DB	1	KKTPEPKKEVTIKANLIYADGKTQTAEFKGTFEETATAEAYADALKKNGEYTVDA 60		
QY	61	DKGYTLNKFAG 72		
DB	61	DKGYTLNKFAG 72		
RESULT 2				
US-08-325-278-1				
; Sequence 1, Application US/08325278				
; Publication No. US20030027283A1				
; GENERAL INFORMATION:				

APPLICANT: Bjvrck, Lars  
APPLICANT: SJvdring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION/DOCKET NUMBER: 450023.401  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
US-08-325-278-1

Query Match 100.0%; Score 371; DB 8; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1e-34;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKTPEEPKEEVTIRKANLIYADGKTQTAEFKGTFFETAEAYRADALKKNGEYTDVA 60  
Db 81 KKTPEEPKEEVTIRKANLIYADGKTQTAEFKGTFFETAEAYRADALKKNGEYTDVA 140  
QY 61 DKGYTLNKFAG 72  
Db 141 DKGYTLNKFAG 152

RESULT 3  
US-08-325-278-3  
Sequence 3, Application US/08325278  
Publication No. US20030027283A1  
GENERAL INFORMATION:  
APPLICANT: Bjvrck, Lars  
APPLICANT: SJvdring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION/DOCKET NUMBER: 450023.401  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
US-08-325-278-3  
Query Match 100.0%; Score 371; DB 8; Length 434;  
Best Local Similarity 100.0%; Pred. No. 1.6e-34;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKTPEEPKEEVTIRKANLIYADGKTQTAEFKGTFFETAEAYRADALKKNGEYTDVA 60  
Db 81 KKTPEEPKEEVTIRKANLIYADGKTQTAEFKGTFFETAEAYRADALKKNGEYTDVA 140  
QY 61 DKGYTLNKFAG 72  
Db 141 DKGYTLNKFAG 152

RESULT 4  
US-09-808-212A-8  
Sequence 8, Application US/09808212A  
Patent No. US20020137918A1  
GENERAL INFORMATION:  
APPLICANT: Gore, Michael Graham  
APPLICANT: Beckingham, Jennifer Ann  
APPLICANT: Roberts, Sian Eleri  
TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
FILE REFERENCE: 100084.41405  
CURRENT APPLICATION NUMBER: US/09/808,212A  
CURRENT FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-8

Query Match 94.9%; Score 352; DB 10; Length 72;  
Best Local Similarity 94.4%; Pred. No. 2.5e-33;  
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KKTPEEPKEEVTIRKANLIYADGKTQTAEFKGTFFETAEAYRADALKKNGEYTDVA 60  
Db 1 KKTPEEPKEEVTIRKANLIYADGKTQTAEFKGTFFETAEAYRADALKKNGEYTDVA 60  
QY 61 DKGYTLNKFAG 72  
Db 61 DKGYTLNKFAG 72

RESULT 5  
US-09-808-212A-10  
Sequence 10, Application US/09808212A  
Patent No. US20020137918A1





```
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-14

Query Match      78.8%; Score 292.5; DB 10; Length 71;
Best Local Similarity 80.6%; Pred. No. 1.7e-26;
Matches 58; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 1 KEKTPPEEVTIKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVA 60
Db 1 KKK-PEEPEEVTIKNLIFADGKTQTAEFGTPEEATAKAYAYADLLAKENGEXTADLE 59
QY 61 DRGYTLNIRFAG 72
Db 60 DGGNTINIRFAG 71

RESULT 10
US-10-345-618-13
; Sequence 13, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:LHL-omp
; OTHER INFORMATION: protein sequence
US-10-345-618-13

Query Match      78.4%; Score 291; DB 12; Length 178;
Best Local Similarity 88.9%; Pred. No. 8e-26;
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVADKGYTLNK 69
Db 12 EEVTKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVADKGYTLNK 71
QY 70 FAG 72
Db 72 FAG 74

RESULT 11
US-10-345-618-8
; Sequence 8, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16

; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:LHL.seq
; OTHER INFORMATION: protein sequence
US-10-345-618-8

Query Match      78.4%; Score 291; DB 12; Length 198;
Best Local Similarity 88.9%; Pred. No. 9.2e-26;
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVADKGYTLNK 69
Db 32 EEVTKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVADKGYTLNK 91
QY 70 FAG 72
Db 92 FAG 94

RESULT 12
US-10-345-618-6
; Sequence 6, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:TLHL protein
; OTHER INFORMATION: sequence
US-10-345-618-6

Query Match      78.4%; Score 291; DB 12; Length 342;
Best Local Similarity 88.9%; Pred. No. 1.8e-25;
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 EEVTKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVADKGYTLNK 69
Db 176 EEVTKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKDNGETYVDVADKGYTLNK 235
QY 70 FAG 72
Db 236 FAG 238

RESULT 13
US-10-345-618-16
; Sequence 16, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
```

; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:ccMTLgI protein  
US-10-345-618-16

Query Match 78.4%; Score 291; DB 12; Length 482;  
Best Local Similarity 88.9%; Pred. No. 2.8e-25;  
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
Qy 10 EETVTKANLIYADGKTTQTAEFGTFEEATAEAYADALKKDNGEYTVDVADKGYTLNIK 69  
Db 323 EETVTKANLIYADGKTTQTAEFGTFEEATAEAYADALKKDNGEYTVDVADKGYTLNIK 382  
Qy 70 FAG 72  
Db 383 FAG 385

RESULT 14  
US-10-345-618-4  
; Sequence 4, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 495  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:CATAB-TEV  
US-10-345-618-4

Query Match 78.4%; Score 291; DB 12; Length 495;  
Best Local Similarity 88.9%; Pred. No. 2.9e-25;  
Matches 56; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
Qy 10 EETVTKANLIYADGKTTQTAEFGTFEEATAEAYADALKKDNGEYTVDVADKGYTLNIK 69  
Db 176 EETVTKANLIYADGKTTQTAEFGTFEEATAEAYADALKKDNGEYTVDVADKGYTLNIK 235  
Qy 70 FAG 72  
Db 236 FAG 238

RESULT 15  
US-09-808-212A-12  
; Sequence 12, Application US/09808212A

; Patent NO. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Elzerl  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 71  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-12

Query Match 78.0%; Score 289.5; DB 10; Length 71;  
Best Local Similarity 80.3%; Pred. No. 3.7e-26;  
Matches 57; Conservative 7; Mismatches 6; Indels 1; Gaps 1;  
Qy 2 EKTPEEPKEEVTIKANLIYADGKTTQTAEFGTFEEATAEAYADALKKDNGEYTVDVAD 61  
Db 1 KETP-EPEEEVTIKANLIYADGKTTQTAEFGTFEEATAEAYADALKKDNGEYTVDVAD 59  
Qy 62 KGYTLNIK FAG 72  
Db 60 KGLTLNIK FAG 70

Search completed: September 3, 2003, 11:42:50  
Job time : 26.6301 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:19:27; Search time 213.041 Seconds  
(without alignments)  
294.149 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_81\_152

Perfect score: 371  
Sequence: 1 KEKTPPEKPEVTKANLIY.....GEYTVADKGYTLNIKFAG 72

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371	100.0	72	23	US-09-808-212A-6
2	371	100.0	305	7	US-08-325-278-1

3	371	100.0	305	7	US-08-325-278A-1	Sequence 1, Appli
4	371	100.0	305	7	US-08-325-278B-1	Sequence 1, Appli
5	371	100.0	367	23	US-09-889-182A-4	Sequence 4, Appli
6	371	100.0	434	7	US-08-325-278-3	Sequence 3, Appli
7	371	100.0	434	7	US-08-325-278A-3	Sequence 3, Appli
8	371	100.0	434	7	US-08-325-278B-3	Sequence 3, Appli
9	371	100.0	467	25	US-09-980-469-12	Sequence 12, Appli
10	371	100.0	719	22	US-09-791-537-10210	Sequence 10210, A
11	371	100.0	719	22	US-09-791-537-96101	Sequence 96101, A
12	352	94.9	72	23	US-09-808-212A-8	Sequence 8, Appli
13	322	86.8	72	23	US-09-808-212A-10	Sequence 10, Appli
14	314	84.6	76	23	US-09-808-212A-4	Sequence 4, Appli
15	314	84.6	78	22	US-09-791-537-22553	Sequence 22553, A
16	308	83.0	75	23	US-09-808-212A-18	Sequence 18, Appli
17	308	83.0	291	8	US-08-446-137A-4	Sequence 4, Appli
18	308	83.0	992	22	US-09-791-537-88366	Sequence 88366, A
19	308	83.0	1027	7	US-08-331-637-2	Sequence 2, Appli
20	308	83.0	1027	8	US-08-446-137A-2	Sequence 2, Appli
21	308	83.0	1027	15	US-09-187-295-2	Sequence 2, Appli
22	299	80.6	182	19	US-09-509-031-2	Sequence 2, Appli
23	299	80.6	182	23	US-09-820-048A-2	Sequence 2, Appli
24	292	78.8	71	23	US-09-808-212A-14	Sequence 14, Appli
25	291	78.4	178	19	US-09-509-031-13	Sequence 13, Appli
26	291	78.4	178	23	US-09-820-048A-13	Sequence 13, Appli
27	291	78.4	198	19	US-09-509-031-8	Sequence 8, Appli
28	291	78.4	198	23	US-09-820-048A-8	Sequence 8, Appli
29	291	78.4	342	19	US-09-509-031-6	Sequence 6, Appli
30	291	78.4	342	23	US-09-820-048A-6	Sequence 6, Appli
31	291	78.4	482	19	US-09-509-031-16	Sequence 16, Appli
32	291	78.4	495	19	US-09-509-031-4	Sequence 4, Appli
33	291	78.4	495	23	US-09-820-048A-4	Sequence 4, Appli
34	289	77.1	71	23	US-09-808-212A-12	Sequence 12, Appli
35	286	77.1	82	23	US-09-808-212A-2	Sequence 2, Appli
36	283	76.3	74	23	US-09-808-212A-16	Sequence 16, Appli
37	90.5	24.4	395	1	PCT-US01-08631-42219	Sequence 42219, A
38	90.5	24.4	395	19	US-09-538-092-857	Sequence 857, App
39	89.5	24.1	463	1	PCT-US01-08631-53631	Sequence 53631, A
40	72.5	19.5	301	27	US-10-179-131-7741	Sequence 7741, Ap
41	72.5	19.5	536	22	US-09-791-537-9081	Sequence 9081, Ap
42	72	19.4	1447	30	US-10-437-963-198658	Sequence 198658,
43	71	19.1	634	22	US-09-791-537-126395	Sequence 126395,
44	69.5	18.7	130	21	US-09-708-427-40518	Sequence 40518, A
45	69.5	18.7	172	21	US-09-708-427-40517	Sequence 40517, A

## ALIGNMENTS

RESULT 1  
US-09-808-212A-6  
; Sequence 6, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808, 212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-6

Query Match 100.0%; Score 371; DB 23; Length 72;  
Best Local Similarity 100.0%; Pred. No. 4.8e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEKPEVTKANLIYADGKTQTAEKFGTFEATAYAYDAALKKDNGEYTVDA 60  
|||||

```
Db 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKDNGEYTVDA 60
QY 61 DRGYTLNIRK 72
| | | | | | | | | |
Db 61 DRGYTLNIRK 72

RESULT 2
US-08-325-278-1
; Sequence 1, Application US/08325278
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; US-08-325-278-1

Query Match 100.0%; Score 371; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 3.5e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKDNGEYTVDA 60
| | | | | | | | | |
Db 81 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKDNGEYTVDA 140
| | | | | | | | | |

QY 61 DRGYTLNIRK 72
| | | | | | | | | |
Db 141 DRGYTLNIRK 152

RESULT 3
US-08-325-278A-1
; Sequence 1, Application US/08325278A
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
```

```
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278A
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-08-325-278A-1

Query Match 100.0%; Score 371; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 3.5e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKDNGEYTVDA 60
| | | | | | | | | |
Db 81 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKDNGEYTVDA 140
| | | | | | | | | |

QY 61 DRGYTLNIRK 72
| | | | | | | | | |
Db 141 DRGYTLNIRK 152

RESULT 4
US-08-325-278B-1
; Sequence 1, Application US/08325278B
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278B
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,278
FILING DATE: 26-OCT-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 450023.401
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 434 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHEetical: NO
ORIGINAL SOURCE:
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055
US-08-325-278-3
Query Match 100.0%; Score 371; DB 7; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.7e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 KETPPEPKPEVTIKANLIYADGKTQTAEFGTTEATAEAYRYADALKKDNGETVDVA 60
Db 81 KETPPEPKPEVTIKANLIYADGKTQTAEFGTTEATAEAYRYADALKKDNGETVDVA 140
QY 61 DKGYTLNKFAG 72
Db 141 DKGYTLNKFAG 152
RESULT 7
US-08-325-278A-3
Sequence 3, Application US/08325278A
GENERAL INFORMATION:
APPLICANT: Bjorck, Lars
Sjoberg, Ulf
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,278A
FILING DATE: 26-Oct-1994
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:

```

```
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-325-278A-3

Query Match 100.0%; Score 371; DB 7; Length 434;
Best Local Similarity 100.0%; Pred. No. 5.7e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEKPEVTIKANLIYADGKTQTAEFKGTFFEATAEAYRYADALKKNGEYTVDA 60
Db 81 KEKTPPEKPEVTIKANLIYADGKTQTAEFKGTFFEATAEAYRYADALKKNGEYTVDA 140
QY 61 DKGYTLNKFAG 72
Db 141 DKGYTLNKFAG 152

RESULT 9
US-09-980-469-12
; Sequence 12, Application US/09980469
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND R
; FILE REFERENCE: 01/22924
; CURRENT APPLICATION NUMBER: US/09/980,469
; CURRENT FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: pUC19-cell1-ProtL-ceXNG-HDEL fusion encoded product
US-09-980-469-12

Query Match 100.0%; Score 371; DB 25; Length 467;
Best Local Similarity 100.0%; Pred. No. 6.3e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEKPEVTIKANLIYADGKTQTAEFKGTFFEATAEAYRYADALKKNGEYTVDA 60
Db 120 KEKTPPEKPEVTIKANLIYADGKTQTAEFKGTFFEATAEAYRYADALKKNGEYTVDA 179
QY 61 DKGYTLNKFAG 72
Db 180 DKGYTLNKFAG 191

RESULT 10
US-09-791-537-10210
; Sequence 10210, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10210
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Peptostreptococcus magnus
US-09-791-537-10210

Query Match 100.0%; Score 371; DB 22; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



Qy	1	KETPEEPKEEVIKANIYADGKTQTAEFKTFFETAAYRADALKKDNGEYTVDA	60
Db	1	KETPEEPKEEVIKANIYADGKTQTAEFKTFFETAAYRYADLLAKENGKYTVDA	60
Qy	61	DKGYYTLNIRFAG	72
Db	61	DKGYYTLNIRFAG	72

RESULT 15  
US-09-791-537-22553  
; Sequence 22553, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomimix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22553  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: pdb 2PTL  
US-09-791-537-22553

Query Match 84.6%; Score 314; DB 22; Length 78;  
Best Local Similarity 81.6%; Pred. No. 3.3e-29;  
Matches 62; Conservative 6; Mismatches 4; Indels 4; Gaps 1;  
QY 1 KEKTPPEP----KEEVTKANLIYADGKQTQTAEFGTPEATAEAYRYADALKKNGEYT 56  
Db 3 KEETPETDSEEEVTKANLIFANGSIQTAEFGTPEKATSEAYAYADTLKKNGEYT 62  
QY 57 VDVADKGYTLNIKFAG 72  
Db 63 VDVADKGYTLNIKFAG 78

Search completed: September 3, 2003, 11:40:07  
Job time : 213.041 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:20:12 ; Search time 10.1096 Seconds  
(without alignments)  
223.701 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_81\_152

Perfect score: 371

Sequence: 1 KEKTEPEKKEVTKANLIY.....GEYTVADVADKGYTLNKFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 148256 seqs, 31410045 residues

Total number of hits satisfying chosen parameters: 148256

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending\_Patents\_AA\_New:\*  
1: /cgn2.6/prodata/1/paa/pct\_NEW\_COMB.pep.\*  
2: /cgn2.6/prodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2.6/prodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2.6/prodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2.6/prodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2.6/prodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2.6/prodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	17.3	291	6	US-10-640-833-4201
2	64	17.3	328	6	US-10-617-320-3318
3	63.5	17.1	914	6	US-10-631-467-789
4	63.5	17.1	1019	1	PCT-US03-21510-62
5	62.5	16.8	157	5	US-09-897-516A-8199
6	62	16.7	930	5	US-09-200-650E-3
7	61	16.4	728	6	US-10-467-534-81
8	61	16.4	977	6	US-10-293-244-1201
9	61	16.4	977	7	US-60-490-890-580
10	61	16.4	1003	6	US-10-293-244-3169
11	60.5	16.3	271	6	US-10-221-131-95
12	60.5	16.3	271	6	US-10-221-131-96
13	60.5	16.3	354	6	US-10-299-636-105
14	60.5	16.3	588	6	US-10-299-636-96
15	60	16.2	247	6	US-10-408-765A-1338
16	60	16.2	1166	5	US-09-200-650E-7
17	59.5	16.0	1849	6	US-10-637-544-2
18	59.5	16.0	2315	6	US-10-603-114-5434
19	59	15.9	102	6	US-10-016-516A-12
20	59	15.9	242	5	US-09-674-548A-244
21	59	15.9	300	5	US-09-397-243D-12
22	59	15.9	300	6	US-10-633-008-10
23	58.5	15.8	383	6	US-10-613-520-1853
24	58.5	15.8	592	1	PCT-US02-21936-30
25	58.5	15.8	634	6	US-10-613-520-1103
26	58.5	15.8	636	6	US-10-613-520-1638

27	58.5	15.8	637	5	US-09-897-516A-7157	Sequence 7157, Ap
28	58.5	15.8	638	6	US-10-613-520-1328	Sequence 1328, Ap
29	58.5	15.8	638	6	US-10-613-520-1341	Sequence 1341, Ap
30	58.5	15.8	638	6	US-10-613-520-1356	Sequence 1356, Ap
31	58.5	15.8	638	6	US-10-613-520-1857	Sequence 1857, Ap
32	58	15.6	1684	6	US-10-408-765A-811	Sequence 811, App
33	57.5	15.5	239	6	US-10-603-113-15852	Sequence 15852, A
34	57.5	15.5	423	7	US-60-491-156-12	Sequence 12, Appl
35	57.5	15.5	489	6	US-10-640-833-4914	Sequence 4914, Ap
36	57.5	15.5	497	6	US-10-617-320-4571	Sequence 4571, Ap
37	57	15.4	1290	6	US-10-617-320-4399	Sequence 4399, Ap
38	57	15.4	1303	6	US-10-640-833-5037	Sequence 5037, Ap
39	57	15.4	1444	1	PCT-US03-20082-11	Sequence 11, Appl
40	57	15.4	1444	6	US-10-609-113-11	Sequence 91, Appl
41	56.5	15.2	136	6	US-10-221-131-91	Sequence 91, Appl
42	56.5	15.2	250	6	US-10-603-113-16721	Sequence 16721, A
43	56.5	15.2	638	6	US-10-631-958-3	Sequence 3, Appl
44	56.5	15.2	682	6	US-10-326-956-523	Sequence 523, App
45	56.5	15.2	713	6	US-10-467-534-97	Sequence 97, Appl

ALIGNMENTS

RESULT 1

US-10-640-833-4201  
; Sequence 4201, Application US/10640833

; GENERAL INFORMATION:

; APPLICANT: Doucette-Stamm, Lynn

; APPLICANT: Bush, David

; APPLICANT: Zeng, Qiangdong

; APPLICANT: Opperman, Timothy

; APPLICANT: Houseweart, Chad Eric

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics

; FILE REFERENCE: PATH03-15

; CURRENT APPLICATION NUMBER: US/10/640,833

; CURRENT FILING DATE: 2003-08-14

; PRIOR APPLICATION NUMBER: US 09/583,110

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 4201

; LENGTH: 291

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-10-640-833-4201

Query Match

Best Local Similarity 17.3%; Score 64; DB 6; Length 291;

Matches 17; Conservative 9; Mismatches 17; Indels 22; Gaps 2;

QY 20 YADGKTQTAEFK-----GTFEATAEAYRYADAKKNGEYTVADVADK 63

Db 165 FIEAKTETETYSQHLGSLVVGSTSVSLMEKLAAY-----KKENPEVITDITNSG 218

QY 64 YTLNI 68

Db 219 SSAGI 223

RESULT 2

US-10-617-320-3318

; Sequence 3318, Application US/10617320

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA



ORGANISM: *Neisseria gonorrhoeae*  
US-10-467-534-81

```

; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 09/523,095
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: JP 2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP 2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: pscM1. MABL1-scfv
US-10-221-131-95.

Query Match 16.3% Score 60.5; DB 6; Length 271;
Best Local Similarity 27.3%; Pred.No.12;
Matches 18; Conservative 11; Mismatches 24; Indels 13; Gaps 2;

QY 20 YADGKTQTAEFGK----TFEATAEAYRYADALKKDN-----GEYTVDVADKGYTL 66
Db 76 YNDGTYKNEFKGKATLTSEKSSAAWELSSLASDESAYVYCARGGYISYDDMGQGTTL 135
QY 67 NIKFAG 72
Db 136 TVSSGG 141

RESULT 12
US-10-221-131-96
; Sequence 96 Application US/10221131
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; TITLE OF INVENTION: POLYPEPTIDE INDUCING APOPTOSIS
; FILE REFERENCE: 065678/0106
; CURRENT APPLICATION NUMBER: US/10/221,131
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 09/523,095
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: JP 2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP 2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: pCHOM1. MABL1-scfv
US-10-221-131-96

Query Match 16.3% Score 60.5; DB 6; Length 271;
Best Local Similarity 27.3%; Pred.No.12;
Matches 18; Conservative 11; Mismatches 24; Indels 13; Gaps 2;

QY 20 YADGKTQTAEFGK----TFEATAEAYRYADALKKDN-----GEYTVDVADKGYTL 66
Db 73 YNDGTYKNEFKGKATLTSEKSSAAWELSSLASDESAYVYCARGGYISYDDMGQGTTL 132
QY 67 NIKFAG 72

```

Db 133 TVSSGG 138

## RESULT 13

US-10-299-636-105

; Sequence 105, Application US/10299636

; GENERAL INFORMATION:

; APPLICANT: Briles, David E

; APPLICANT: McDaniel, Larry S

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Crain, Marilyn J

; APPLICANT: Hollingshead, Susan

; APPLICANT: Tart, Rebecca

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF

; FILE REFERENCE: 57909/361

; CURRENT APPLICATION NUMBER: US/10/299,636

; PRIOR FILING DATE: 2002-11-19

; PRIOR APPLICATION NUMBER: 08/714,741

; PRIOR FILING DATE: 1996-09-16

; PRIOR APPLICATION NUMBER: 08/529,055

; PRIOR FILING DATE: 1995-09-15

; NUMBER OF SEQ ID NOS: 111

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 105

; LENGTH: 354

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-10-299-636-105

Query Match

Best Local Similarity 16.3%; Score 60.5; DB 6; Length 354;

Matches 21; Conservative 11; Mismatches 32; Indels 9; Gaps 1;

## Query

Db 119 KQAPELTKLEAKLEAEKATEAKOKVDAEVAPOAKTAELENQVHRLEQELKEI 178

## Query

Db 52 NGEYTVDVADKGY 64

## Db

179 DESESEDYAKEGF 191

## RESULT 14

US-10-299-636-96

; Sequence 96, Application US/10299636

; GENERAL INFORMATION:

; APPLICANT: Briles, David E

; APPLICANT: McDaniel, Larry S

; APPLICANT: Swiatlo, Edwin

; APPLICANT: Yother, Janet

; APPLICANT: Crain, Marilyn J

; APPLICANT: Hollingshead, Susan

; APPLICANT: Tart, Rebecca

; APPLICANT: Brooks-Walter, Alexis

; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF

; FILE REFERENCE: 57909/361

; CURRENT APPLICATION NUMBER: US/10/299,636

; PRIOR FILING DATE: 2002-11-19

; PRIOR APPLICATION NUMBER: 08/714,741

; PRIOR FILING DATE: 1996-09-16

; PRIOR APPLICATION NUMBER: 08/529,055

; PRIOR FILING DATE: 1995-09-15

; NUMBER OF SEQ ID NOS: 111

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 96

; LENGTH: 588

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-10-299-636-96

Query Match

Best Local Similarity 16.3%; Score 60.5; DB 6; Length 588;

Best Local Similarity 28.8%; Pred. No. 30;  
Matches 21; Conservative 11; Mismatches 32; Indels 9; Gaps 1;

## Query

Db 136 KQAPELTKLEAKLEAEKATEAKOKVDAEVAPOAKTAELENQVHRLEQELKEI 195

## Query

Db 52 NGEYTVDVADKGY 64

## Db

196 DESESEDYAKEGF 208

## RESULT 15

US-10-408-765A-1338

; Sequence 1338, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1338

; LENGTH: 247

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-1338

Query Match

Best Local Similarity 16.2%; Score 60; DB 6; Length 247;

Matches 18; Conservative 11; Mismatches 20; Indels 4; Gaps 2;

## Query

Db 65 EKTPEQIMQEKQIEAKI--EDLENEIEEVKVAFEIKKALDRMLSTALKKN 115

Search completed: September 3, 2003, 11:40:55

Job time : 10.1096 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:14:37 ; Search time 17.0137 Seconds  
(without alignments)  
406.975 Million cell updates/sec

Title: US-08-325-278b-1-copy\_81\_152  
Perfect score: 371  
Sequence: 1 KEKTEPEKKEVTIKANLIY.....GEYTVDVADKGYTLNIKIFAG 72

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	371	100.0	455	2 A45063	immunoglobulin-bin
2	371	100.0	719	2 A42808	Ig light chain-bin
3	308	83.0	992	2 S54396	protein L precursor
4	171.5	46.2	74	2 A34483	Ig light chain-bin
5	90.5	24.4	395	2 A43700	BNS1 protein - hum
6	89.5	24.1	1576	2 AE0249	probable hemolysin
7	69.5	18.7	256	2 S54803	merozoite surface
8	69	18.6	529	2 S62194	hypothetical prote
9	68.5	18.5	243	2 F72719	hypothetical prote
10	68.5	18.5	279	2 S67132	hypothetical prote
11	68.5	18.5	280	2 I55577	Fc gamma (IgG) rec
12	68.5	18.5	344	2 A41357	Fc gamma (IgG) rec
13	68.5	18.5	374	1 A39878	Fc gamma (IgG) rec
14	66.5	17.9	160	2 I47163	cytolytic trigger
15	66	17.8	448	2 H84782	late embryogenesis
16	65.5	17.7	718	2 F90512	hypothetical prote
17	65.5	17.7	829	2 E64114	translation initia
18	65.5	17.7	458	2 C82199	RTX toxin RtxA Vcl
19	64	17.3	291	2 G95243	hypothetical prote
20	64	17.3	312	2 D98108	hypothetical prote
21	64	17.3	448	2 JC6171	late embryogenesis
22	64	17.3	929	2 T51932	kinesin (imported)
23	63.5	17.1	255	2 F83769	hypothetical prote
24	63.5	17.1	436	2 T18638	hypothetical prote
25	63.5	17.1	582	2 S42613	membrane protein p
26	63.5	17.1	796	2 S50672	hypothetical prote
27	63	17.0	401	2 AH0364	probable HlyD fami
28	63	17.0	499	2 T10680	cytochrome P450 mo
29	63	17.0	950	2 B87611	TonB-dependent rec

30	62.5	16.8	280	2 S35103	bone sialoprotein
31	62.5	16.8	451	1 D64424	tlGD homolog MJ099
32	62.5	16.8	476	1 WZEC	tryptophanase (EC
33	62.5	16.8	476	2 E91209	tryptophanase [imp
34	62.5	16.8	476	2 H86055	tryptophanase [imp
35	62.5	16.8	487	2 B71873	hypothetical prote
36	62.5	16.8	1179	2 T35093	DNA-directed DNA p
37	62	16.7	266	2 A12289	hypothetical prote
38	62	16.7	384	2 G86287	hypothetical prote
39	62	16.7	455	2 D81284	probable fumarate
40	62	16.7	516	2 B64902	probable oligopept
41	62	16.7	658	2 S37494	squalene-hopane cy
42	62	16.7	871	2 D86355	protein T16E15.12
43	61.5	16.6	334	2 C84117	cell shape determi
44	61.5	16.6	479	2 T47561	late embryogenesis
45	61.5	16.6	528	2 D83874	nickel ABC transpo

ALIGNMENTS

RESULT 1  
A45063  
immunoglobulin-binding protein LG - Peptostreptococcus magnus  
C:Species: Peptostreptococcus magnus  
C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999  
C:Accession: A45063  
R:Kihlberg, B.M.; Sjoerding, U.; Kastern, W.; Bjorck, L.  
J. Biol. Chem. 267, 25583-25588, 1992  
A:Title: Protein LG: a hybrid molecule with unique immunoglobulin binding properties.  
A:Reference number: A45063; MUID:93094283; PMID:1460053  
A:Accession: A45063  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-455 <KIR>  
A:Cross-references: GB:S50809; NID:q261705; PIDN:AAA03280-1; PID:q261706  
A>Note: sequence extracted from NCBI backbone (NCBIN:120302, NCBI:120303)

Query Match 100.0%; Score 371; DB 2; Length 455;  
Best Local Similarity 100.0%; Pred. No. 6.6e-30;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KEKTEPEKKEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKNGEYTVDA 60
DB	102	KEKTEPEKKEVTIKANLIYADGKTQTAEFKGTFFETAETAEAYRYADALKKNGEYTVDA 161
QY	61	DKGYTLNIKIFAG 72
DB	162	DKGYTLNIKIFAG 173

RESULT 2  
A42808  
Ig light chain-binding protein precursor - Peptostreptococcus magnus  
N:Alternate names: protein L  
C:Species: Peptostreptococcus magnus  
C>Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 07-May-1999  
C:Accession: A42808; A41493  
R:Kastern, W.; Sjoerding, U.; Bjorck, L.  
J. Biol. Chem. 267, 12820-12825, 1992  
A:Title: Structure of peptostreptococcal protein L and identification of a repeated i  
A:Reference number: A42808; MUID:92316971; PMID:1618782  
A:Accession: A42808  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-719 <KAS>  
A:Cross-references: GB:M86697  
R:Kastern, W.; Holst, E.; Nielsen, E.; Sjoerding, U.; Bjorck, L.  
Infect. Immun. 58, 1217-1222, 1990  
A:Title: Protein L, a bacterial immunoglobulin-binding protein and possible virulence  
A:Reference number: A41493; MUID:90215984; PMID:2108927  
A:Accession: A41493  
A:Status: preliminary

A;Molecule type: DNA  
A;Residues: 202-275 <KA2>  
C;Keywords: immunoglobulin

Query Match	100.0%;	Score 371;	DB 2;	Length 719;
Best Local Similarity	100.0%;	Pred. No. 1.1e-29;		
Matches 72;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	KKTPPEPKKEEVTIKANLIYADGKTQTAEFGTFEEATAEAYRYADALKKDNGEYTVDVA	60	
Dd	174	KKTPPEPKKEEVTIKANLIYADGKTQTAEFGTFEEATAEAYRYADALKKDNGEYTVDVA	233	
QY	61	DKGYTLNIKFKAG	72	
Dd	234	DKGYTLNIKFKAG	245	

RESULT 3  
S54396  
protein L precursor - Peptostreptococcus magnus (strain 3316)  
C:Species: Peptostreptococcus magnus  
A:Variety: strain 3316  
C:date: 27-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 15-Oct-1999  
C:Accession: S54396  
R:Murphy, J.P.; Duggleby, C.J.; Atkinson, M.A.; Trowern, A.R.; Atkinson, T.; Goward, C.F.  
Mol. Microbiol. 12, 911-920, 1994  
A:title: The functional units of a peptostreptococcal protein L.  
A:Reference number: S54396; MUID:95020613; PMID:7934898  
A:Accession: S54396  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-992 <MUR>  
A:Cross-references: EMBL:L04466; NID:g150673; PIDN:AAA67503.1; PID:g150674

```

Query Match      83.0%; Score 308; DB 2; Length 992;
Best Local Similarity 84.3%; Pred. No. 3.5e-23;
Matches 59; Conservative 4; Mismatches 7; Indels 0; Gaps 0

QY 3 KYPEPKBEVTIKANLIYADGKTOTAEFKGTFEATAEAYRYADALKKONGEYTVDVADK 62
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 468 EPEPKBEVTIKNLIYFADGKTOTAEFKGTFEATAEAYRYADLLAKYNGEYTDLEDG 527
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 63 GYTLNLIKFAG 72
      ||||| |||||
Db 528 GYTLNLIKFAG 537
      ||||| |||||

```

RESULT 4  
A34483  
Ig light chain-binding protein L - Peptostreptococcus magnus (fragments)  
C:Species: Peptostreptococcus magnus  
C:date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 16-Aug-1996  
C:Accession: A34483  
R:Akerstroem, B.; Bjoerck, L.  
J. Biol. Chem. 264, 19740-19746, 1989  
A:title: Protein L: an immunoglobulin light chain-binding bacterial protein. Characterization of the protein structure and its interaction with immunoglobulin light chains  
A:Reference number: A34483; MUID:90062074; PMID:2479638  
A:Accession: A34483  
A:Molecule type: protein  
A:Residues: 1-74 <AKE>  
C:Keywords: immunoglobulin

```

Query Match      46.28; Score 171.5; DB 2; Length 74;
Best Local Similarity 58.28; Pred. No. 1.3e-10;
Matches 39; Conservative 5; Mismatches 4; Indels 19; Gaps 2;

QY 2 EKTPEEPKEEVTIKANLIYADGKTQTAEFG-----TPEETAAYRYA 45
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 ETTPEEPKEEVTIKANLIYADGKTETAEFGPEETPEKPEVDGYAEEETAAA---K 65
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY	46 DALKKDN	52
	:      :	
Db	66 EALKNDP	72

RESULT 5

BN51 protein - human  
A43700

C:Species: Homo sapiens (man)  
C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 21-Jul-2000

C:Accession: A43700

E:Ittmann, M.; Greco, A.; Basillico, C.  
Mol. Cell. Biol. 7, 3386-3393, 1987

A:Title: Isolation of the human gene that complements a temperature-sensitive cell cy

A:Reference number: A43700; MUID:88065472; PMID:3683386

A:Accession: A43700

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-395 <177>

A:Cross-references: GB:M17754; NID:g179512; PIDN:AAA51838.1; PID:g179513

C:Genetics:

A:Gene: GDB:BN51t

A:Cross-references: GDB:119728; OMIM:187280

A:Map position: 8pter-8q24

```

Query Match      24.4%; Score 90.5; DB 2; Length 395;
Best Local Similarity 33.3%; Pred. No. 0.13;
Matches 27; Conservative 10; Mismatches 25; Indels 19; Gaps 3;

QY 3 KTEEPKEEVTIIKANLIYD-----GKTQTAEFKGTTEETAERYADALUKK 50
  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 KKEEPEEVTVKKEKRRDRDRQREGHGRGRRRPEVIQSHSFEQGPAEMMK-----KK 117
                                     : : : | | : | | |
QY 51 DNGEYTVDVADKG--YTLNLIK 69
                                     : | | | | | : | | |
Db 118 GNWDKTVDSVDMGSPSHIINIK 138

```

RESULT 6  
AE0249.  
Probable hemolysin YPO2045 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AE0249  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentiss,  
delo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dou-  
lin, M.; Rutherford, K.; Simmonds, T.; Skelton, J.; Stevens, K.; Whitehead, S.;  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Title: Genbank number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AE0249  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1576 <RUR>  
A:Cross-references: GB:AL590842; PIDN:CAC90857.1; PID:g15980056; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO2045

```

Query Match      24.1%; Score 89.5; DB 2; Length 1576;
Best Local Similarity 27.5%; Pred. No. 0.7;
Matches 30; Conservative 11; Mismatches 25; Indels 43; Gaps 4;

Qy 1 KETTPPEPKKEEVTIKANLIYADGKTQTQAPKPGTF-----EE----- 36
   | | | | | | | | | | | | | | | | | | | | | |
Db 1001 KANTTQEKGEVSLRG-----GWTATQETKGLGVKAETSQGDSYAEEMLVGNINAKSG 1054

Qy 37 -----ATAEAYRYADALKKDNGEVTVD-----VADKGYTINIRFAG 72
   | | | | | | | | | | | | | | | | | | | | | |
Db 1055 VSIKTTGDAYYYATINTEGGNGDITDAGNNLYDFDQVDSORSNNIKPSFG 1103

```

RESULT 7  
S54803  
merozoite surface glycoprotein precursor - *Theileria parva* (fragment)  
C;Species: *Theileria parva*  
C;date: 08-Jul-1995 #sequence revision 21-Jul-1995 #text change 09-Sep-1997

Query Match	18.5%;	Score 68.5;	DB 2;	Length 243;	
Best Local Similarity	28.4%;	Pred. No. 12;			
Matches	19;	Conservative 14;	Mismatches 33;	Indels 1;	Gaps 1;
QY	6	EEPEEVTIKANLIYADCKTQTAFKGTFFEATAEAYRYADALKDKNGEYTVDVADKGYT	65		
Db	134	DAPLDYYSVKADSSGVKRSKPKGGYVYVADSSGVKLATA-PATRGYETWIDVDKSSG	192		
QY	66	LNKIFAG	72		
Db	193	VTIEFEG	199		
RESULT 10					
S67132					
hypothetical protein YOR239w - yeast (Saccharomycetes cerevisiae)					
N:Alternate names: hypothetical protein O5240					
C:Species: saccharomycetes cerevisiae					
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002					
C:Accession: S67132					
R:Boyer, J.; Fairhead, C.; Gaillon, L.; Galisson, F.; Michaux, G.; Thierry, A.					
submitted to the Protein Sequence Database, July 1996					
A:Reference number: S67104					
A:Accession: S67132					
A:Molecule type: DNA					
A:Residues: 1-279 <BOY>					
A:Cross-references: EMBL:Z75147; NID:g1420546; PID:e252406; PID:g1420547; GSPD:3					
A:Experimental source: strain S288C					
C:Genetics:					
A:Gene: SGD:ABP140; MIPS:YOR239w					
A:Cross-references: SGD:S0005765					
A:Map position: 15R					
Query Match	18.5%;	Score 68.5;	DB 2;	Length 279;	
Best Local Similarity	27.5%;	Pred. No. 14;			
Matches	19;	Conservative 14;	Mismatches 27;	Indels 9;	Gaps 1;
QY	1	KEKTEPPKERVITKANLIYADGKT-----QTAEFKGTFFEATAEAYRYADALKKD	51		
Db	81	KPETNEEDDESGSMENKIYSGENADINVNDFQYKEMENTGAEVLASSVEESDAIQEG	140		
QY	52	NGEYTVDNA	60		
Db	141	VAEETEGIA	149		
RESULT 11					
I55577					
Fc gamma (IgG) receptor I-B splice form 1 precursor - human					
N:Alternate names: CD64					
N:Contains: Fc-gamma (IgG) receptor I-B splice form 2					
C:Species: Homo sapiens (man)					
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000					
C:Accession: I55577; I70303					
R:Porces, A.J.; Redecha, P.B.; Doebele, R.; Pan, L.C.; Salmon, J.E.; Kimberly, J. Clin. Invest. 90, 2102-2109, 1992					
A:Title: Novel Fc gamma receptor I family gene products in human mononuclear					
A:Reference number: I55577; MUID:93055454; PMID:1430234					
A:Accession: I55577					
A>Status: preliminary; translated from GB/EMBL/DBDJ					
A:Molecule type: mRNA					
A:Residues: 1-280 <RES>					
A:Cross-references: GB:L03419; NID:g182460; PIDN:AAA35825.1; PID:g292023					
A:Note: splice form B1					
A:Accession: I70303					
A>Status: preliminary; translated from GB/EMBL/DBDJ					
A:Molecule type: mRNA					
A:Residues: 1-10,103-153 'A',155-280 <RE2>					
A:Cross-references: GB:L03420; NID:g182461; PIDN:AAA35826.1; PID:g292024					
A:Experimental source: mononuclear cells					
A:Note: splice form B2					
C:Comment: This receptor does not bind monomeric IgG with high affinity.					
C:Genetics:					

A;Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A;Residues: 1-374 <RES>  
A;Cross-references: GB:L03418; NID:g184840; PIDN:AAA36049.1; PID:g292169  
R:Allen, J.M.; Seed, B.  
Science 243, 378-381, 1989  
A;Title: Isolation and expression of functional high-affinity Fc receptor complementary DNA  
A;Reference number: A41357; MUID:89100284; PMID:2911749  
A;Accession: B41357  
A:Molecule type: mRNA  
A;Residues: 1-24,'S','26-337','T','339-374 <ALL>  
A;Cross-references: GB:X14356; GB:M21091; NID:g31331; PIDN:CAA32537.1; PID:g31332  
R:Allen, J.M.; Seed, B.  
Nucleic Acids Res. 16, 11824, 1988  
A;Title: Nucleotide sequence of three cDNAs for the human high affinity Fc receptor (FcγRIII)  
A;Reference number: S03018; MUID:89098339; PMID:2974947  
A;Accession: S03018  
A:Molecule type: mRNA  
A;Residues: 1-24,'S','26-337','T','339-374 <ALL>  
A;Cross-references: EMBL:X14356; NID:g31331; PIDN:CAA32537.1; PID:g31332  
R:Perez, C.; Wietzerbin, J.; Benech, P.D.  
Mol. Cell. Biol. 13, 2182-2192, 1993  
A;Title: Two cis-DNA elements involved in myeloid-cell-specific expression and gamma chain enhancer activity  
A;Reference number: I57525; MUID:93204964; PMID:84555606  
A;Accession: I57525  
A;Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A;Residues: 1-8 <RES>  
A;Cross-references: GB:S57204; NID:g298692; PIDN:AAD13887.1; PID:g4261587  
C:Genetics:  
A;Gene: GDB:FCGR1A; CD64  
A;Cross-references: GDB:I35911; OMIM:146760  
A;Map position: lq21-lq21  
C:Superfamily: Fc gamma receptor I; immunoglobulin homology  
C;Keywords: alternative splicing; glycoprotein; immunoglobulin receptor; transmembrane protein  
F;1-15/Domain: signal sequence #status predicted <SIG>  
F;16-292/Domain: extracellular #status predicted <EXT>  
F;117-170/Domain: immunoglobulin homology <IMM2>  
F;293-313/Domain: transmembrane #status predicted <TM>  
F;59,78,152,159,163,195,240/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.5%; Score 68.5; DB 1; Length 374;  
Best Local Similarity 31.2%; Pred.No. 20;  
Matches 15; Conservative 10; Mismatches 22; Indels 1; Gaps 1;

QY 9 KEETVTKANLIYAGKQTQAEF-KGTPEEATAEAYRYADALKKDNGEY 55  
:|||:: :: :|||:: :|||:: :|||:: :|||::  
Db 36 EETVTLLHCEVLHPGSSSTQWFLNGATQTSPTSYRTSASVNDSGEY 83  
:|||:: :: :|||:: :|||:: :|||:: :|||::

RESULT 14  
I47163  
cytolytic trigger molecule G7 - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C;Accession: I47163  
R:Halloran, P.J.; Sweeney, S.E.; Strohmaier, C.M.; Kim, Y.B.  
J. Immunol. 153, 2631-2641, 1994  
A;Title: Molecular cloning and identification of the porcine cytolytic trigger molecule cDNA  
A;Reference number: I47163; MUID:94358430; PMID:8077673  
A;Accession: I47163  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A;Residues: 1-160 <HAL>  
A;Cross-references: EMBL:U08991; NID:g476167; PIDN:AAA57188.1; PID:g476168  
C:Superfamily: Fc gamma receptor III; immunoglobulin homology  
F;14-65/Domain: immunoglobulin homology <IM>

Query Match 17.9%; Score 66.5; DB 2; Length 160;  
Best Local Similarity 34.7%; Pred.No. 13;  
Matches 17; Conservative 6; Mismatches 25; Indels 1; Gaps 1;

Search completed: September 3, 2003, 11:24:29  
Job time : 18.0137 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:52 ; Search time 9.13229 Seconds  
(without alignments)  
371.130 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_81\_152

Perfect score: 371

Sequence: 1 KEKTEPEKPEVIRKANLIY.....GEYTVDVADKGYTLNIFKAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.5	26.5	398	1	RPC4_HUMAN
2	93.5	25.2	398	1	RPC4_MOUSE
3	72.5	19.5	536	1	INR2_SHEEP
4	71	19.1	634	1	GYR2_BORHE
5	68.5	18.5	374	1	FCGI_HUMAN
6	68.5	18.5	627	1	ABPX_YEAST
7	65.5	17.7	829	1	IF2_HAEIN
8	64.5	17.4	370	1	MYOM_APLCA
9	64	17.3	600	1	DNAK_MYCHY
10	63.5	17.1	796	1	YE29_YEAST
11	63.5	17.1	1461	1	A10B_HUMAN
12	63	17.0	499	1	C831_ARATH
13	62.5	16.8	300	1	SIAL_PIG
14	62.5	16.8	451	1	Y996_METJA
15	62.5	16.8	471	1	TNAA_ECO57
16	62.5	16.8	471	1	TNAA_ECOLI
17	62.5	16.8	642	1	DNAK_FRATU
18	62.5	16.8	1179	1	DP3A_STRCO
19	62	16.7	516	1	YD05_ECOLI
20	62	16.7	658	1	SQRC_2YMO
21	61	16.4	1410	1	RBL1_HUMAN
22	61	16.4	1433	1	SUBF_BACSU
23	60.5	16.3	444	1	SLAP_LACAC
24	60.5	16.3	637	1	DNAK_PSEAE
25	60.5	16.3	1454	1	CSP2_HUMAN
26	60.5	16.3	2805	1	MAPA_HUMAN
27	60	16.2	499	1	LEUL_PYRFU
28	60	16.2	459	1	SYR_XANAC
29	60	16.2	1829	1	Y296_HUMAN
30	59.5	16.0	269	1	T2S1_STRFI
31	59.5	16.0	350	1	Y957_METJA
32	59.5	16.0	404	1	ASSV_VIBCH
33	59.5	16.0	553	1	HIS5_EWENI

## ALIGNMENTS

### RESULT 1

RPC4_HUMAN	1176	1	SLAP_BACSH
ID	RPC4_HUMAN	STANDARD;	PRT; 398 AA.
AC	P05423; Q9BPV7; Q9BP21; Q9BXB3;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	15-SEP-2003 (Rel. 42, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	DNA-directed RNA polymerase III 47 kDa polypeptide (EC 2.7.7.6) (RNA		
DE	polymerase C subunit 4) (RPC4) (RPC53) (BN51 protein).		
GN	BN51T OR BN51 OR POLR3D.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=88065472; PubMed=3683386;		
RA	Iltmann M., Greco A., Basilico C.;		
RT	"Isolation of the human gene that complements a temperature-sensitive		
RL	cell cycle mutation in BHK cells.";		
RL	Mol. Cell. Biol. 7:3386-3393(1987).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22278866; PubMed=12391170;		
RA	Hu P., Wu S., Sun Y., Yuan C.-C., Kobayashi R., Myers M.P.,		
RA	Hernandez N.;		
RT	"Characterization of human RNA polymerase III identifies orthologues		
RT	for Saccharomyces cerevisiae RNA polymerase III subunits.";		
RL	Mol. Cell. Biol. 22:8044-8055(2002).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung, and Lymph;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,		
RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length		
RT	human and mouse cDNA sequences.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[4]		
RP	SEQUENCE OF 1-50 FROM N.A.		
RX	SEQUENCE=21282972; PubMed=11279001;		

P38537 bacillus sp  
P32386 saccharomyc  
P07891 chlamydomon  
Q9pab5 ureaplasma  
Q88792 mus musculus  
Q8u089 pyrococcus  
Q56232 thermus the  
Q934g3 aquifex pyr  
Q28942 sus scrofa  
Q8rc30 thermoanaer  
Q28938 sus scrofa

34 59.5 16.0 1176 1 SLAP\_BACSH  
35 59.5 16.0 1661 1 YBT1\_YEAST  
36 59 15.9 140 1 ATPE\_CHLRE  
37 59 15.9 242 1 Y376\_UREPA  
38 59 15.9 300 1 JAM1\_MOUSE  
39 59 15.9 324 1 TRPD\_PYRFU  
40 59 15.9 385 1 AAT\_THETH  
41 59 15.9 871 1 SAA\_AQUPY  
42 58.5 15.8 236 1 KDGW\_ERWCH  
43 58.5 15.8 257 1 FCG3\_PIG  
44 58.5 15.8 261 1 RNPB\_THETN  
45 58.5 15.8 324 1 IL2B\_PIG

RA Chong S.S., Hu P., Hernandez N.;  
 RT "Reconstitution of transcription from the human U6 small nuclear RNA  
 RT promoter with eight recombinant polypeptides and a partially purified  
 RT RNA polymerase III complex.";  
 RL J. Biol. Chem. 276:20727-20734(2001).  
 RN [5]  
 RP SEQUENCE OF 1-12 FROM N.A.  
 RA Ittmann M.;  
 RT "Promoter structure and cell cycle control of the B51 cell cycle  
 RT gene, which encodes a subunit of RNA polymerase III.";  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
 CC of DNA into RNA using the four ribonucleoside triphosphates as  
 CC substrates. Complements a temperature-sensitive cell cycle  
 CC mutation in BHK cells.  
 CC [CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA)(N).  
 CC -!- SUBUNIT: RNA polymerase III consists of about 15 different  
 CC subunits. Interacts with RPC5.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- CAUTION: Ref.3 (AAH03039 and AAH00516) sequence differs from that  
 CC shown due to a frameshift in position 175.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M17754; AAA51838.1; -;  
 DR EMBL; AY092086; AAM18216.1; -;  
 DR EMBL; BC002603; AAH02603.1; -;  
 DR EMBL; BC003039; AAH03039.1; ALT\_FRAME.  
 DR EMBL; BC004484; AAH04484.1; -;  
 DR EMBL; BC000516; AAH00516.1; ALT\_FRAME.  
 DR EMBL; AF346574; AAK15371.1; -;  
 DR EMBL; L15301; AAY23377.1; -;  
 DR PIR; A43700; A43700.  
 DR Genew; HGNC:1080; BNS1T.  
 DR MIM; 187280; -;  
 DR GO; GO:000074; P:regulation of cell cycle; TAS.  
 DR Pfam; PF05132; RNA\_pol\_Rpc4; 1.  
 DR Transferrase; DNA-directed RNA polymerase; Transcription;  
 KW Nuclear protein.  
 KW CONFLICT 9 18 EPSTPGGPRP -> RPARQGPDL (IN REF. 1).  
 FT CONFLICT 9 12 EPST -> RPAR (IN REF. 5).  
 FT CONFLICT 26 40 LIGRRPAPLTPGRL -> SSGGGGLPSPPAV (IN REF.  
 FT 1).  
 FT CONFLICT 98 98 G -> R (IN REF. 1).  
 FT CONFLICT 233 233 K -> R (IN REF. 2).  
 FT CONFLICT 286 286 P -> L (IN REF. 2).  
 SQ SEQUENCE 398 AA; 44396 MW; CD8AFF3257B78410 CRC64;  
 Query Match 26.5%; Score 98.5; DB 1; Length 398;  
 Best Local Similarity 34.6%; Pred. No. 0.0066;  
 Matches 28; Conservative 10; Mismatches 24; Indels 19; Gaps 3;  
 QY 3 KTEPEEKVEVTKANLIYAD-----GKTQTAEFGTGFEBATAEAYRYADLRK 50  
 Db | ||||| | | | : : : || : ||  
 66 KIKEPKEEVTKREKRRDRDRREGHGRGRPEVIOHSIFEGPAEMMK-----KK 120  
 QY 51 DNGEYTVVDKGG--YTLNIK 69  
 | : ||||| : : : |||  
 Db 121 GNWDKTVDSVDMGSPSHIINIK 141  
 RESULT 2  
 RPC4\_MOUSE  
 ID RPC4\_MOUSE STANDARD; PRT; 398 AA.  
 AC Q91WD1; Q9CZ02;  
 DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE DNA-directed RNA polymerase III 47 kDa polypeptide (EC 2.7.7.6) (RNA  
 DE polymerase C subunit 4) (RPC4).  
 GN BNS1T OR POLR3D.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton K., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
 CC of DNA into RNA using the four ribonucleoside triphosphates as  
 CC substrates (By similarity).  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA)(N).  
 CC -!- SUBUNIT: RNA polymerase III consists of about 15 different  
 CC subunits. Interacts with RPC5 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----



FT	DOMAIN	268	536	CYTOPLASMIC (POTENTIAL).	3 X 5 AA TANDEM REPEATS OF S-L-E-D-C.
FT	DOMAIN	358	372	1.	
FT	REPEAT	358	362	2.	
FT	REPEAT	363	367	3.	
FT	REPEAT	368	372	BY SIMILARITY.	
FT	DISULFID	85	93	BY SIMILARITY.	
FT	DISULFID	210	230	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	58	58	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	101	101	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	147	147	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE	536 AA;	60260 MW;	1059F27D80545150 CRC64;	

Query Match 19.5%; Score 72.5; DB 1; Length 536;  
Best Local Similarity 31.1%; Pred. No. 4.2;  
Matches 23; Conservative 12; Mismatches 30; Indels 9; Gaps 4;

QY	4	TPEPKEVITKANLIYADQKQTQTAEPKGT---	FEEATAEAYRYADAL--	KKONGEYTV	58
DB	461	SPSPPEETAIVLEEDLS	TESSESLIVASEG	TQLPFTDPSMECLRPQDALS	DKDTS
QY	59	VADKGY--	TLNIK	69	
DB	521	IGD-GYIVRQVNLK	533		

RESULT 4

ID	GYRB_BORHE	STANDARD;	PRT;	634 AA.
AC	Q9ZFK1;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	DNA gyrase subunit B (EC 5.99.1.3).			
GN	GYRB.			
OS	Borrelia hermslii.			
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID=140;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HSL;			
RT	Samuels D.S., Kimmel B.J., Huang W.M.;			
RA	"Mutations in Borrelia hermslii gyrb confer resistance to coumermycin			
RL	Al.";			
RT	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-			
CC	STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE			
CC	INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED			
CC	DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.			
CC	-1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining			
CC	of double-stranded DNA.			
CC	-1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA			
CC	BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE			
CC	ENZYME FORMS AN A2B2 TETRAMER.			
CC	-1- SIMILARITY: Belongs to the type II topoisomerase family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AF098862; AAC72846.1;			
DR	HSP; P06982; IAJ6.			
DR	InterPro; IPR003594; ATPbind_Atptase.			
DR	InterPro; IPR002288; DNA_gyraseB_C.			
DR	InterPro; IPR000565; DNA_gyrb.			
DR	InterPro; IPR001241; DNA_topoisolI.			
DR	InterPro; IPR006171; Toprim_dom.			
DR	Pfam; PF00204; DNA_gyraseB.1.			

```

DR Pfam; PF00986; DNA_gyraseB_C; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF01751; Toprim; 1.
DR PROSITE; PR00418; TPI2FAMILY.
DR ProDom; PD149633; DNA_gyraseB_C; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00433; TOP2c; 1.
DR TIGRFAMs; TIGR01059; gyfB; 1.
DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
DR Topoisomerase; Isomerase; ATP-binding.
SQ SEQUENCE 634 AA; 70848 MW; 5E58C5F32126833C CRC64;

Query Match
Best Local Similarity 19.1%; Score 71; DB 1; Length 634;
Matches 20; Conservative 10; Mismatches 23; Indels 2; Gaps 1;

QY 17 NLIYADGKTQTAEFGTPEATAEAYRVADALKDNGEYTVDAVKGYT--LNLIK 69
   : : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 270 NINTREGTHVAGFKSGFLKANSEAFROSKISKDVPSLTLDKFEGLTAVISIK 324

RESULT 5
FCG1_HUMAN STANDARD; PRT; 374 AA.
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE High affinity immunoglobulin gamma Fc receptor I precursor (Fc-gamma
DE RI) (FCRI) (IGG Fc receptor I) (CD64).
GN FCGRIA OR FCGRI OR FCGI OR IGFRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89098339; PubMed=2974947;
RA Allen J.M., Seed B.;
RT "Nucleotide sequence of three cDNAs for the human high affinity Fc
RT receptor (FcRI).";
RL Nucleic Acids Res. 16:11824-11824(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89100284; PubMed=2911749;
RA Allen J.M., Seed B.;
RT "Isolation and expression of functional high-affinity Fc receptor
RT complementary DNAs.";
RL Science 243:378-381(1989).
CC 1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS GAMMA. HIGH
CC 1- AFFINITY RECEPTOR.
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=P12314-1; Sequence=Displayed;
CC
CC IsoId=P12314-2; Sequence=VSP_002637;
CC 1- TISSUE SPECIFICITY: MONOCYTE/MACROPHAGE-SPECIFIC.
CC 1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC 1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC 1- DATABASE: NAME=PROW; NOTE=CD guide CD64 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd64.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; X14356; CAA32537.1; -.

```

```

DR EMBL; X14355; CAA32536.1; -.
DR HSP; P12319; IALS.
DR Genew; HGNC:3613; FCGRIA.
DR MIM; 146760; -.
DR GO; GO:0005057; F:receptor signaling protein activity; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0006911; P:phagocytosis, engulfment; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR Pfam; PF00047; Ig; 3.
DR Pfam; SM00409; Ig; 3.
DR PROSITE; PS00835; IG-LIKE; 3.
KW IgG-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
KW Immunoglobulin domain; Repeat; Alternative splicing; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 374
FT HIGH AFFINITY IMMUNOGLOBULIN GAMMA FC
FT RECEPTOR I.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT IG-LIKE C2-TYPE 1.
FT IG-LIKE C2-TYPE 2.
FT IG-LIKE C2-TYPE 3.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT HEKVTSSLSQEDRHLEELKCEQKEQLQEGVHRKEPOGA
FT T -> GOALEAPTOGCA (in isoform B).
FT /FTID-VSP_002637.
FT L -> T.
FT /FTID-VAR_003953.
FT N -> V.
FT /FTID-VAR_003954.
FT S -> T (IN REF. 1; CAA32536).
FT CONFLICT 25 25
FT SEQUENCE 374 AA; 42605 MW; 2C2AA8103ECF16E6 CRC64;
SQ
Query Match 18.5%; Score 68.5; DB 1; Length 374;
Best Local Similarity 31.2%; Pred. No. 7.4;
Matches 15; Conservative 10; Mismatches 22; Indels 1; Gaps 1;

QY 9 KEETVTKANLIYADGKTQTAEF-KCTFEATAEAYRVADALKDNGEY 55
   : | | | : : : | | | : : : | | | : : : | | |
Db 36 EETVTLHCEVLHPGSSSQWFLNGTATQTSTPSYRITSASVDSGEY 83

RESULT 6
ABPX_YEAST STANDARD; PRT; 627 AA.
AC Q08641; Q08644;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin-binding protein ABP140.
GN ABP140 OR YOR239W/YOR240W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RX MEDLINE=97127829; PubMed=8972580;
RA Boyer J., Michaux G., Fairhead C., Gaillon L., DuJon B.;
RT "Sequence and analysis of a 26.9 kb fragment from chromosome XV of

```

```

RT the yeast Saccharomyces cerevisiae.";
RL Yeast 12:1575-1586(1996).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., SEQUENCE OF 1-12; 38-72; 102-115; 244-286;
RP 560-574 AND 598-608, FUNCTION, AND SUBCELLULAR LOCATION.
RC STRAIN-BJ5457;
RX MEDLINE=98127445; PubMed=9467951;
RA Asakura T., Sasaki T., Nagano F., Satoh A., Obaishi H., Nishioka H.,
RA Imamura H., Hotta K., Tanaka K., Nakanishi H., Takai Y.;
RT "Isolation and characterization of a novel actin filament-binding
RT protein from Saccharomyces cerevisiae.";
RL Oncogene 16:121-130(1998).
CC -!- FUNCTION: Binds F-actin and shows weak F-actin crosslinking
CC activity.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and cortical cytoskeleton.
CC -!- MISCELLANEOUS: A RIBOSOMAL FRAMESHIFT OCCURS BETWEEN THE CODONS
CC FOR LEU-276 AND GLY-277.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z75147; CAA99460.1; ALT-SEQ.
CC EMBL: Z75147; CAA99461.1; ALT-SEQ.
CC SGD: S0005765; ABP140.
CC DR GO: 0005884; C:actin filament; IDA.
CC DR GO: 0003780; F:actin cross-linking activity; IDA.
CC DR GO: 0003036; P:actin cytoskeleton organization and biogenesis; IDA.
CC InterPro: IPR001601; Methyltransf.
CC DR InterPro: IPR000051; SAM_bind.
CC Actin-binding; Cytoskeleton; Ribosomal frameshift.
FT INIT_MET 0
FT SEQUENCE 627 AA; 71354 MW; D4E55F9485412F39 CRC64;
Query Match 18.5%; Score 68.5; DB 1; Length 627;
Best Local Similarity 27.5%; Pred. No. 13;
Matches 19; Conservative 14; Mismatches 27; Indels 9; Gaps 1;
QY 1 KEKTPPEKPEEYTIKANLIYADGKT-----QTAEPKGTGFEATAEAYADALKD 51
Db 80 KPETNEDEEFGSMSENKIYSGENADINVDQYKEMTGNTGAELVSSVEESDAIORG 139
QY 52 NGEYTVDDVA 60
Db 140 VAEETEGIA 148
RESULT 7
IF2_HAEIN ID IF2_HAEIN STANDARD; PRT; 829 AA.
AC P44323;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Translation initiation factor IF-2.
GN INF OR H11284.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

```

---

```

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: One of the essential components for the initiation of
CC protein synthesis. Protects formylmethionyl-tRNA from spontaneous
CC hydrolysis and promotes its binding to the 30S ribosomal subunits.
CC Also involved in the hydrolysis of GTP during the formation of the
CC 70S ribosomal complex.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32808; AAC22933.1; -.
CC PIR: E64114; E64114.
CC TIGR: H11284; -.
CC HAMAP: MF_00100; -.
CC DR InterPro: IPR000795; EF_GTPbind.
CC DR InterPro: IPR004161; EFTU_D2.
CC DR InterPro: IPR00178; IF2.
CC DR InterPro: IPR006847; IF2_N.
CC DR InterPro: IPR005225; Small_GTP.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 2.
CC Pfam: PF04760; IF2_N; 1.
CC ProDom: PD186100; IF2; 1.
CC TIGRFAMs: TIGR00487; IF-2; 1.
CC TIGRFAMs: TIGR00231; Small_GTP; 1.
CC PROSITE: PS01176; IF2; 1.
CC DR Initiation factor; Protein biosynthesis; GTP-binding;
CC KW Complete proteome.
FT DOMAIN 331 480 G-DOMAIN.
FT NP_BIND 337 344 GTP (BY SIMILARITY).
FT NP_BIND 384 388 GTP (BY SIMILARITY).
FT NP_BIND 438 441 GTP (BY SIMILARITY).
SQ SEQUENCE 829 AA; 90551 MW; 75B20AC4CF610AF7 CRC64;
Query Match 17.7%; Score 65.5; DB 1; Length 829;
Best Local Similarity 29.1%; Pred. No. 35;
Matches 25; Conservative 12; Mismatches 32; Indels 17; Gaps 3;
QY 1 KEKTPPEKPEEYTIKANLIYAD---GKTQTAEPKGTGTFEE-----ATAEAYRYA 45
Db 86 KAEATAPKVSADVSRKAKSVESKEKRGKAGEALRRKAEELAKQAEQARRAVEAKRYA 145
QY 46 DALKKDNGEYTVDDVADKGYTLNPKFA 71
Db 146 EADSDNESSESDYSD--YNLSRYA 169
RESULT 8
MYON_APLCA ID MYON_APLCA STANDARD; PRT; 370 AA.
AC P15513; Q07974; Q27916;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myomodulin neuropeptides precursor [Contains: Myomodulin A (MM-A)
DE (PMNSMLRL-amide) (Neuron B16 peptide); Myomodulin B (MM-B) (GSYRMRL-
DE amide); Myomodulin D (MM-D) (GLSMLRL-amide); Myomodulin F (MM-F)
DE (SLNMLRL-amide); Myomodulin G (MM-G) (TLSMLRL-amide); Myomodulin H
DE (MM-H) (GLHMLRL-amide); Myomodulin I (MM-I) (SLSMLRL-amide)].
DE MYOMODI.
GN MYOMODI.

```

OS Aplysia californica (California sea hare).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Agastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidae;  
 OC Aplousiidae; Aplousiidae; Aplysia.  
 OX NCBI\_TaxID=6500;  
 RN [1]  
 RC SEQUENCE FROM N.A., VARIANT LEU-362, AND TISSUE SPECIFICITY.  
 RP TISSUE=Abdominal ganglion, Cerebral ganglion, and CNS;  
 RX MEDLINE=93135828; PubMed=8422272;  
 RA Lopez V., Wickham L., Desgroselliers L.;  
 RT "Molecular cloning of myomodulin cDNA, a neuropeptide precursor gene  
 RL expressed in neuron L10 of Aplysia californica.";  
 RN DNA Cell Biol. 12:53-61(1993).  
 RN [2]  
 RC SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RP TISSUE=Buccal ganglion;  
 RX MEDLINE=93340709; PubMed=8340812;  
 RA Miller M.W., Reushausen S., Vittek A., Stamm S., Kupfermann I.,  
 RA Brosius J., Weiss K.R.;  
 RT "The myomodulin-related neuropeptides: characterization of a gene  
 RL encoding a family of peptide cotransmitters in Aplysia.";  
 RN J. Neurosci. 13:3358-3367(1993).  
 RN [3]  
 RC SEQUENCE OF 53-60. AMIDATION, FUNCTION, AND TISSUE SPECIFICITY.  
 RP TISSUE=Buccal muscle;  
 RX MEDLINE=92158798; PubMed=1788132;  
 RA Cropper E.C., Vilim F.S., Alevisos A., Tenenbaum R., Kolks M.A.G.,  
 RA Rosen S., Kupfermann I., Weiss K.R.;  
 RT "Structure, bioactivity, and cellular localization of myomodulin B: a  
 RL novel Aplysia peptide.";  
 RN Peptides 12:683-690(1991).  
 RN [4]  
 RC SEQUENCE OF 203-209 AND 358-364, AND FUNCTION.  
 RP TISSUE=Buccal muscle;  
 RX MEDLINE=96063267; PubMed=7472354;  
 RA Brezina V., Bank B., Cropper E.C., Rosen S., Vilim F.S.,  
 RA Kupfermann I., Weiss K.R.;  
 RT "Nine members of the myomodulin family of peptide cotransmitters at  
 RL the B16-ARC neuromuscular junction of Aplysia.";  
 RN J. Neurophysiol. 74:54-72(1995).  
 RN [5]  
 RC SEQUENCE OF MYOMODULIN A, AMIDATION, AND FUNCTION.  
 RP TISSUE=Buccal muscle;  
 RX MEDLINE=87261010; PubMed=3474664;  
 RA Cropper E.C., Tenenbaum R., Kolks M.A.G., Kupfermann I., Weiss K.R.;  
 RT "Myomodulin: a bioactive neuropeptide present in an identified  
 RL cholinergic buccal motor neuron of Aplysia.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 84:5483-5486(1987).  
 CC -!- FUNCTION: Exogenous application of myomodulins potentiates ARC  
 CC muscle contraction.  
 CC -!- TISSUE SPECIFICITY: Expressed in all ganglia of the CNS, but only  
 CC in a subset of neurons including L10 in the abdominal ganglion and  
 CC B16 in the buccal ganglion.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L01421; AAA27758.1; -  
 CC EMBL: S55210; AAB25131.1; -  
 CC EMBL: S55211; AAB25132.1; -  
 CC EMBL: S64300; AAB27697.1; -  
 CC Neuropeptide; Amidation; Repeat; Cleavage on pair of basic residues;  
 CC Signal.  
 CC 1 18 POTENTIAL.  
 CC SIGNAL  
 CC 19 50 POTENTIAL.  
 CC PROPEP  
 CC 53 60 MYOMODULIN B.  
 CC PEPTIDE  
 CC 63 69 MYOMODULIN H (POTENTIAL).  
 CC PROPEP  
 CC 73 190 POTENTIAL.

FT PEPTIDE 193  
 FT PEPTIDE 203  
 FT PROPEP 213  
 FT PEPTIDE 240  
 FT PEPTIDE 250  
 FT PEPTIDE 260  
 FT PEPTIDE 270  
 FT PEPTIDE 280  
 FT PEPTIDE 290  
 FT PEPTIDE 300  
 FT PEPTIDE 310  
 FT PEPTIDE 320  
 FT PEPTIDE 330  
 FT PROPEP 340  
 FT PROPEP 350  
 FT PROPEP 358  
 FT MOD\_RES 60  
 FT MOD\_RES 69  
 FT MOD\_RES 199  
 FT MOD\_RES 209  
 FT MOD\_RES 246  
 FT MOD\_RES 256  
 FT MOD\_RES 266  
 FT MOD\_RES 276  
 FT MOD\_RES 286  
 FT MOD\_RES 296  
 FT MOD\_RES 306  
 FT MOD\_RES 316  
 FT MOD\_RES 326  
 FT MOD\_RES 336  
 FT MOD\_RES 346  
 FT MOD\_RES 364  
 FT VARIANT 362  
 FT CONFLICT 231  
 SQ SEQUENCE 370 AA; 42253 MW; 3A792085939C88CB CRC64;  
 Query Match 17.4%; Score 64.5; DB 1; Length 370;  
 Best Local Similarity 29.0%; Pred. No. 19;  
 Matches 27; Conservative 13; Mismatches 24; Indels 29; Gaps 6;  
 QY 1 KEKTEPEPEKEVTIKR--NLI---YADGKTQTAEFKGT-----FEEATAEAY-RY----- 44  
 Db 71 KRGGPVEPESEENLETLNLLQGYSDVPEPSEFDDTDLAYPYEYDAPAHPRERSTP 130  
 QY 45 -----ADALKK-----DNGETVDVADKGY 64  
 Db 131 PTDGVVAPDVLQKGSSEFEDFGDSQDSDEGY 163  
 RESULT 9  
 ID DNAAK\_MYCHY STANDARD; PRT; 600 AA.  
 AC Q49539;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chaperone protein dnak (Heat shock protein 70) (Heat shock 70 kDa  
 DE protein) (HSP70) (65 kDa protein) (P65).  
 GN DNAAK.  
 OS Mycoplasma hyopneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=232;  
 RA Chou S.Y., Shiu D.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.



FT	DOMAIN	1274	1291	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1292	1316	POTENTIAL.
FT	DOMAIN	1317	1461	CYTOPLASMIC (POTENTIAL).
FT	MOD_RES	433	433	PHOSPHORYLATION (BY SIMILARITY).
FT	METAL	1055	1055	MAGNESIUM (BY SIMILARITY).
FT	METAL	1059	1059	MAGNESIUM (BY SIMILARITY).
FT	VARSPLIC	1	136	MAI5YDSSKHHWRQWVGGPHCPSEFTPLLSPKPGQSVN
FT				LTOQRVPNNNSIFHQDWEEVSRRYPGNRTCTTKYILFTFL
FT				PRNLFQPHRWANLYFLVLVILNMPMSMEVPHREITMLPLA
FT				IVLFLVIMIKDGMDFKRRFDKALNCNSIRYE -> MKKE
FT				GKRWKRRDKRVVSNLLPEGWSHKENPNHRHGNVQIKT
FT				SKYTVLSVPKNEIQOLHRFANLYFVGIAVLNFIPIVNAFO
FT				PEVSMPICTVLAVTAIKDAWEDLRRYKSDKVINNRECLY
FT				S (in isoform C).
FT				/FTId-VSP_007305.
FT				AKRLSTPELSDSDMEETQYCLFSARWAOQDPATMRSQKG
FT	VARSPLIC	461	529	AQPLRRSQARVPDQGHYRQSRMGHRES -> GIEAPKGS
FT				PLSKRQIPALLRNEIKDILLALLEAVWHFHKLLPVSLWSS
FT				LSQIRAVPICTKLSFVYKG (in isoform B and
FT				isoform C).
FT				/FTId-VSP_007306.
FT	VARSPLIC	530	1461	Missing (in isoform B and isoform C).
FT				/FTId-VSP_007307.
FT	CONFLICT	217	217	C -> R (IN REF. 2).
FT	CONFLICT	234	234	F -> S (IN REF. 1).
FT	SEQUENCE	1461 AA;	165390 MW;	2676B90416B8541 CRC64;
Qy	Query Match	17.1%;	Score 63.5;	DB 1; Length 1461;
	Best Local Similarity	40.5%;	Pred. No. 1e+02;	Indels 1; Gaps
	Matches 17;	Conservative 5;	Mismatches 19;	
Db	3 KTFPEEKEEVTIKANLIYADGKTOAEFGTFEATAEAYRY 44			
	981 KTPSITSEAVVPEAGLV-IDGKTLNAIFOGKLEKKFLLELTQY 1021			
RESULT 12				
C831_ARATH				
ID	C831_ARATH	STANDARD;	PRT;	499 AA.
AC	O65782;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DE	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome P450 83B1 (EC 1.14.-.-).			
GN	CYP83B1 OR AT4G31500 OR F3L17.70.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi			
OX	NCBI_TaxID=3702;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia; TISSUE=Seedling;			
RC	MEDLINE=98281573; PubMed=9620263;			
RA	Mizutani M., Ward E., Ohta D.;			
RT	"Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of			
RT	cDNAs, differential expression, and RFLP mapping of multiple			
RL	cytochromes P450";			
RL	Plant Mol. Biol. 37:39-52(1998).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RC	MEDLINE=20083488; PubMed=10617198;			
RA	Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,			
RA	Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terryn N.,			
RA	Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,			
RA	Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,			
RA	Kreis M., Delsený M., Puigdomenech P., Watson M., Schmidheini T.,			
RA	Reichert B., Portecelle D., Perez-Alonso M., Bontury M., Bancroft I.,			
RA	Vos P., Hehseisel J., Zimmermann W., Wedler H., Ridley P.,			
RA	Langham S.-A., McCullagh B., Bilham L., Robben J.,			
RA	Van der Schueren J., Grymonpre B., Chuang Y.-J., Vandenbussche F.,			
RA	Braeken M., Welltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,			

RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,  
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,  
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T., H.,  
RA Dose S., de Haan M., Maarle A.C., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,  
RA Neumann S., Agrillon A., Vitale D., Liguori R., Piravandi E.,  
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Feiber R.,  
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,  
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
RA Nelson J., Spith J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M., Martienssen R., McCombie W.R.,  
RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RA thaliana.";  
RA Nature 402:769-777(1999).  
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D78598; BAA3531.1; -;  
DR EMBL; AL080283; CAB45909.1; -;  
DR EMBL; AL161579; CAB79868.1; -;  
DR PIR; T10680; T10680.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.  
FT TRANSMEM 3 23 POTENTIAL.  
FT METAL 441 441 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
SQ SEQUENCE 499 AA; 56846 MW; CCA2D733E7D00D7B CRC64;  
Query Match 17.0%; Score 63; DB 1; Length 499;  
Best Local Similarity 26.8%; Pred. No. 37;  
Matches 22; Conservative 7; Mismatches 31; Indels 22; Gaps 2;  
QY 5 PEPKPEVTIKANLIYADGKTQTAEFKGTPEEA-----TAEA-----Y 42  
Db 257 PNPKEETESFIDLLMQIKQDFSKFTHENVKAMLDIVPGTDAARAVVWMTYLI 316  
QY 43 RYADALKKNGEYTVADVADKY 64  
Db 317 KYPEAMKKAQDEVRISVIGDKGY 338  
RESULT 13

SIAL\_PIG STANDARD; .PRT; 300 AA.  
AC P31936;  
DT 01-JUN-1993 (Rel. 26, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Bone sialoprotein II (BSP II) (Cell-binding sialoprotein) (Integrin-  
DE binding sialoprotein).  
GN IBSP.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Theria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
[1]  
RN RP SEQUENCE OF 1-20.  
RC TISSUE=Bone;  
RX MEDLINE=90237064; PubMed=2332443;  
RA Zhang Q., Domenicucci C., Goldberg H.A., Wrana J.L., Sodek J.;  
RT "Characterization of fetal porcine bone sialoproteins, secreted  
RT phosphoprotein I (SPPI, osteopontin), bone sialoprotein, and a 23-kDa  
RT glycoprotein. Demonstration that the 23-kDa glycoprotein is derived  
RT from the carboxyl terminus of SPPI.";  
RL J. Biol. Chem. 265:7583-7589(1990).  
[2]  
RN RP SEQUENCE OF 21-300 FROM N.A.  
RX MEDLINE=94142682; PubMed=8309422;  
RA Shapiro H.S., Chen J., Wrana J.L., Zhang Q., Blum M., Sodek J.;  
RT "Characterization of porcine bone sialoprotein: primary structure and  
RT cellular expression.";  
RL Matrix 13:431-440(1993).  
CC -1- FUNCTION: BINDS TIGHTLY TO HYDROXYAPATITE. APPEARS TO FORM AN  
CC INTEGRAL PART OF THE MINERALIZED MATRIX. PROBABLY IMPORTANT TO  
CC CELL-MATRIX INTERACTION. PROMOTES ARG-GLY-ASP-DEPENDENT CELL  
CC ATTACHMENT.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MISCELLANEOUS: IT IS POSSIBLE THAT THE SEGMENTS OF CLUSTERED  
CC CARBOXYL GROUPS MEDIATE THE STRONG BINDING TO HYDROXYAPATITE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L10363; AAA19822.1; -;  
DR PIR; S40032; S35103.  
KW Glycoprotein; Sialic acid; Biominalization; Cell adhesion.  
FT DOMAIN 49 160 GLU-RICH (ACIDIC).  
FT DOMAIN 61 68 POLY-GLU.  
FT DOMAIN 141 151 POLY-GLU.  
FT SITE 272 274 CELL ATTACHMENT SITE.  
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 300 AA; 33026 MW; C853E8EBDE156B6 CRC64;  
Query Match 16.8%; Score 62.5; DB 1; Length 300;  
Best Local Similarity 28.8%; Pred. No. 24;  
Matches 17; Conservative 10; Mismatches 21; Indels 11; Gaps 3;  
QY 5 PEPKPEVTIKANLIYADGKTQTAEFKGTPEEAAYRAYADALKKNGEYTVADVADKG 63  
Db 222 PTTPQD-ISGTTLPSPGKTTTPEYGEYEQTAHEY-----DNG-YEYESENG 269  
RESULT 14  
ID Y996\_METJA STANDARD; .PRT; 451 AA.  
AC Q58403;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)



```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0996.
GN MJ0996.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Botodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE TLDD/PMBA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67542; AAB99001.1; -.
CC TIGR: MJ0996; -.
CC InterPro: IPR002510; Pmba_TlDD.
CC Pfam: PF01523; Pmba_TlDD; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 451 AA; 49813 MW; D5245F2E9551FB79 CRC64;
Query Match 16.8%; Score 62.5; DB 1; Length 451;
Best Local Similarity 24.7%; Pred. No. 37;
Matches 22; Conservative 13; Mismatches 25; Indels 29; Gaps 4;
QY 2 EKTPEPKKEVTKANLIYADKGTQAEKPKFEEATAYAYADALKDNGEYF-----56
Db 352 EELLEDTKEGIFLKS--RGQVDTG--KGIFQFSAVEAYLI-----ENGELTQVLKD 400
QY 57 -----VDVADKGYTLNKFAG 72
Db 401 AGLSGELDLFKVDVATKDFELSVGYCG 429
RESULT 15
TNAAL_ECO57
ID TNAAL_ECO57 STANDARD; PRT; 471 AA.
AC O8XB34;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).
GN TNAAL OR Z5203 OR ECS4645.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Groetbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
```

```
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -!- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +
CC NH(3).
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Tryptophan catabolism.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE005602; AAG58908.1; ALT_INIT.
CC EMBL: AF002566; BAB38068.1; ALT_INIT.
CC HAMAP: MF_00544; -.
CC InterPro: IPR001597; Beta_elim_lyase.
CC Pfam: PF01212; Beta_elim_lyase; 1.
CC ProDom: PD005927; Beta_elim_lyase; 1.
CC PROSITE: PS00853; BETA_ELIM_LYASE; 1.
KW Tryptophan catabolism; Lyase; Pyridoxal phosphate; Complete proteome.
FT BINDING 270 270 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 471 AA; 52789 MW; 49DD2D41BD9D0034 CRC64;
Query Match 16.8%; Score 62.5; DB 1; Length 471;
Best Local Similarity 48.5%; Pred. No. 39;
Matches 16; Conservative 4; Mismatches 8; Indels 5; Gaps 2;
QY 24 KQTQAEFKG-TFEEATAEAYRYADAL-----KKD 51
Db 239 KQREAEYKDWITIEQITRETYKYADMLAMSAKKD 271
Search completed: September 3, 2003, 11:20:04
Job time : 11.1233 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:17 ; Search time 48.3288 Seconds  
(without alignments)  
236.470 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_225\_296

Perfect score: 370

Sequence: 1 KEKTPPEPKKEVTKANLIY.....GKYTADLEGGYTNIRFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370	100.0	72	21	AAV82540 Peptostreptococcus
2	370	100.0	305	14	AA42993 Immunoglobulin lig
3	370	100.0	434	14	AA42994 Sequence encoding
4	370	100.0	467	22	AAAB31372 Amino acid sequenc
5	339	91.6	72	21	AAV82538 Peptostreptococcus
6	329	88.9	75	21	AAV82544 Peptostreptococcus
7	329	88.9	291	14	AA42204 Immunoglobulin bin
8	329	88.9	1027	14	AA42203 Protein L. Peptoc
9	329	88.9	1027	14	AA43699 Protein L. Peptoc

10	322	87.0	72	21	AAV82538
11	322	87.0	367	21	AA10432
12	315.5	85.3	71	21	AAV82542
13	306	82.7	74	21	AAV82543
14	304	82.2	82	21	AAV82536
15	297	80.3	82	21	AAV82545
16	297	80.3	82	21	AAV82548
17	295	79.7	82	21	AAV82546
18	291	78.6	82	21	AAV82547
19	269	72.7	76	21	AAV82537
20	254	68.6	182	18	AAW32479
21	254	68.6	182	20	AAV06907
22	251.5	68.0	71	21	AAV82541
23	246	66.5	178	18	AAW32486
24	246	66.5	178	20	AAV06913
25	246	66.5	198	18	AAW32481
26	246	66.5	198	20	AAV06910
27	246	66.5	342	18	AAW32482
28	246	66.5	342	20	AAV06909
29	246	66.5	482	20	AAV06915
30	246	66.5	495	18	AAW32480
31	246	66.5	495	20	AAV06908
32	69.5	18.8	2062	23	ABB91373
33	68.5	18.5	1045	23	ABB91371
34	67.5	18.2	368	22	ABB61686
35	67.5	18.2	836	23	AAAG68288
36	67.5	18.2	871	23	AAAG68285
37	66.5	18.0	1032	23	ABB91372
38	65	17.6	87	22	ABG51447
39	65	17.6	87	22	ABB36598
40	65	17.6	87	22	ABB21937
41	65	17.6	87	22	AAW57361
42	65	17.6	87	22	AAW69757
43	65	17.6	87	22	AAW17577
44	65	17.6	87	23	ABG39384
45	65	17.6	156	20	AAV36276

#### ALIGNMENTS

RESULT 1  
AAV82540  
ID AAV82540 standard; Protein; 72 AA.

XX  
AC AAV82540;

XX  
XX

XX  
DT 20-JUL-2000 (first entry)

XX  
DE Peptostreptococcus strain 312 protein L domain B4 protein sequence.

XX  
KW Immunoglobulin light chain binding protein; PpL; protein L;

XX  
KW Peptostreptococcus; human immunoglobulin kappa chain;

XX  
KW immunoaffinity chromatography.

XX  
OS Peptostreptococcus sp.

XX  
PN WO200015803-A1.

XX  
PD 23-MAR-2000.

XX  
PF 14-SEP-1999; 99WO-GB03048.

XX  
PR 14-SEP-1998; 98GB-0019998.

XX  
PR 26-APR-1999; 99GB-0009578.

XX  
(ACTI-) ACTINOVA LTD.

XX  
Gore MG, Beckingham JA, Roberts SE;

XX  
WPI: 2000-271441/23.

XX  
N-PSDB; AAA08429.

XX

PT New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
PT more at pH8 with respect to human immunoglobulin kappa-chain -  
XX  
XX  
PS Disclosure; Page 44-45; 56pp; English.  
XX  
XX The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (PI) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
XX Sequence 72 AA;  
Query Match 100.0%; Score 370; DB 21; Length 72;  
Best Local Similarity 100.0%; Pred. No. 6.9e-39;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDLE 60  
Db 1 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDLE 60  
QY 61 DGGYTINIRFAG 72  
Db 61 DGGYTINIRFAG 72  
RESULT 2  
AAR42993  
ID AAR42993 standard; Protein; 305 AA.  
XX  
XX AAR42993;  
DT 25-MAR-2003 (updated)  
DT 16-MAY-1994 (first entry)  
XX  
XX Immunoglobulin light chain binding protein (Protein L).  
DE Immunoglobulin; light chain; binding; identification; purification;  
KW separation.  
KW  
XX  
XX E. coli LE392/pHDL, DSM 7054.  
XX  
XX Key Location/Qualifiers  
FH Domain 5..305  
FT /label= B1 immunoglobulin light chain binding  
FT Domain 81..305  
FT /label= B2 immunoglobulin light chain binding  
FT Domain 153..305  
FT /label= B3 immunoglobulin light chain binding  
FT Domain 225..305  
FT /label= B4 immunoglobulin light chain binding  
FT Domain 297..305  
FT /label= B5 immunoglobulin light chain binding  
XX WO9322342-A1.  
XX  
XX 11-NOV-1993.  
XX  
XX 28-APR-1993; 93WO-SE00375.  
XX  
XX 28-APR-1992; 92SE-0001331.  
XX  
XX (HIGH-) HIGHTECH RECEPTOR AB.  
XX  
XX Bjoerck L, Sjoerbring U;

XX  
DR WPI; 1993-368722/46.  
DR N-PSDB; AAQ50452.  
XX  
XX New protein L binding light chains of all immunoglobulin classes  
PT - for binding purifying and identifying immunoglobulin, also  
PT related DNA, vectors and host cells  
XX  
XX Claim 1; Page 36; 71pp; English.  
XX  
XX The protein (Protein L) is capable of binding to immunoglobulin G  
CC light chains. It is useful for binding, separating (purifying) and  
CC identifying immunoglobulin and for removing immunoglobulin molecules  
CC from serum. Hybrid proteins of the L protein can bind all human  
CC immunoglobulin classes and many immunoglobulins from other species.  
CC They are highly soluble and retain their binding activity at high  
CC temperatures over a pH range of 3-10. They can be immobilised  
CC without loss of activity.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 305 AA;  
Query Match 100.0%; Score 370; DB 14; Length 305;  
Best Local Similarity 100.0%; Pred. No. 4.8e-38;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDLE 60  
Db 225 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDLE 284  
QY 61 DGGYTINIRFAG 72  
Db 285 DGGYTINIRFAG 296  
RESULT 3  
AAR42994  
ID AAR42994 standard; Protein; 434 AA.  
XX  
XX AAR42994;  
DT 25-MAR-2003 (updated)  
DT 16-MAY-1994 (first entry)  
XX  
XX Sequence encoding immunoglobulin light chain binding protein.  
DE Immunoglobulin; light chain; binding; identification; purification;  
KW separation; ss.  
KW  
XX  
XX E. coli L392/pHDLG, DSM 7055.  
XX  
XX Key Location/Qualifiers  
FH Domain 5..305  
FT /label= B1 immunoglobulin light chain binding  
FT Domain 81..305  
FT /label= B2 immunoglobulin light chain binding  
FT Domain 153..305  
FT /label= B3 immunoglobulin light chain binding  
FT Domain 225..305  
FT /label= B4 immunoglobulin light chain binding  
FT Domain 297..305  
FT /label= B5 immunoglobulin light chain binding  
XX 309..434  
XX /label= C1 immunoglobulin heavy chain binding  
XX 364..434  
XX /label= D intermediate immunoglobulin heavy  
XX chain binding domain.

```

FT Domain 379..434
FT /label= C2 immunoglobulin heavy chain binding
FT domain.
XX WO9322342-A1.
XX 11-NOV-1993.
XX 28-APR-1993; 93WO-S500375.
XX 28-APR-1992; 92SE-0001331.
XX (HIGH-) HIGHTECH RECEPTOR AB.
XX Bjoerck L, Sjoerbring U;
XX WPI: 1993-368722/46.
XX P-PSDB; AAR42994.
XX New protein L binding light chains of all immunoglobulin classes
XX - for binding purifying and identifying immunoglobulin, also
XX related DNA, vectors and host cells
XX Claim 6; Page 39-40; 7lpp; English.
XX Protein L (AAR42993) is capable of binding to immunoglobulin G light
XX chains. It is useful for binding, separating (purifying) and
XX identifying immunoglobulin and for removing immunoglobulin molecules
XX from serum. This is the coding sequence of one hybrid protein of
XX the L protein. The hybrid proteins can bind all human
XX immunoglobulin classes and many immunoglobulins from other species.
XX They are highly soluble and retain their binding activity at high
XX temperatures over a pH range of 3-10. They can be immobilised
XX without loss of activity.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 434 AA;
XX
Query Match 100.0%; Score 370; DB 14; Length 434;
Best Local Similarity 100.0%; Pred. No. 7.7e-38;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KETPEPEKPEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60
DB 225 KETPEPEKPEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 284
QY 61 DGGYTTINIRFAG 72
DB 285 DGGYTTINIRFAG 296
XX
RESULT 4
AAB31372
ID AAB31372 standard; Protein; 467 AA.
XX
AC AAB31372;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of protein L/CBD cex/ER retaining peptide fusion.
XX
KW Protein production; food processing; protein antibiotic; feed enzyme;
KW protein L: CBD cex protein; cell signal peptide.
XX
OS Synthetic.
XX
PN WO20007174-A1.
XX
PD 21-DEC-2000.
XX
PF 07-JUN-2000; 2000WO-IL00330.
XX
PR 10-JUN-1999; 99US-0329234.

```

```

XX
PA (CBDT-) CBD TECHNOLOGIES LTD.
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Shani Z, Shoseyov O;
XX
DR WPI: 2001-112219/12.
DR N-PSDB; AAF24730.
XX
PT Expressing and isolating recombinant protein in a plant, useful for
PT producing large quantities of recombinant proteins, by expressing a
PT fusion protein including a cellulose binding peptide fused to a
PT recombinant protein
XX
XX Example; Fig 2a; 87pp; English.
XX
CC The specification describes a method for expressing and isolating a
CC recombinant protein in a plant. The method comprising expressing a
CC fusion protein including the recombinant protein and a cellulose
CC binding peptide fused to it, where the fusion protein is
CC compartmentalised and sequestered within plant cells, plant derived
CC tissue or cultured plant cells. The method is useful for obtaining large
CC quantities of the recombinant proteins and protein products in a simple
CC and cost-effective manner. Recombinant proteins may be used commercially,
CC such as in the food processing industry, e.g. glucoamylases and glucose
CC isomerases are used for converting starch to high fructose corn syrup,
CC proteinases for the hydrolysis of high molecular weight proteins and in
CC manufacturing leather or alcoholic beverages, pectinesterases for
CC pectin hydrolysis in food industry, lipases for cleaving ester linkage
CC in triglycerides, and for effluent treatment. The recombinant proteins
CC may further be used to produce protein antibiotics, which can be used
CC in healing processes, and to produce animal feed enzymes. The present
CC sequence represents a fusion protein of the invention, and comprises a
CC fusion of a cell signal peptide, protein L, CBD cex, and an endoplasmic
CC reticulum retaining peptide.
XX
SQ Sequence 467 AA;
XX
Query Match 100.0%; Score 370; DB 22; Length 467;
Best Local Similarity 100.0%; Pred. No. 8.5e-38;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KETPEPEKPEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60
DB 265 KETPEPEKPEVITKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 324
QY 61 DGGYTTINIRFAG 72
DB 325 DGGYTTINIRFAG 336
XX
RESULT 5
AAY82539
ID AAY82539 standard; Protein; 72 AA.
XX
AC AAY82539;
XX
DT 20-JUL-2000 (first entry)
XX
DE Peptostreptococcus strain 312 protein L domain B3 protein sequence.
XX
KW Immunoglobulin light chain binding protein; PpL; protein L;
KW Peptostreptococcus; human immunoglobulin kappa chain;
KW immunoaffinity chromatography.
XX
OS Peptostreptococcus sp.
XX
PN WO200015803-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-GB03048.
XX

```

PR 14-SEP-1998; 98GB-0019998.  
 PR 26-APR-1999; 99GB-0009578.  
 XX  
 PA (ACTI-) ACTINOVA LTD.  
 XX  
 XX Gore MG, Beckingham JA, Roberts SE;  
 PI  
 XX WPI; 2000-271441/23.  
 DR N-PSDB; AAA08428.  
 DR  
 XX New modified immunoglobulin light chain binding protein, useful in  
 PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
 PT more at pH8 with respect to human immunoglobulin kappa-chain -  
 XX  
 XX Disclosure; Page 43; 56pp; English.  
 PS  
 XX The present invention describes an immunoglobulin (Ig) light chain  
 CC binding protein (Pl) which has been modified by one or more amino acid  
 CC substitutions such that the dissociation constant (Kd) of the protein  
 CC with respect to human Ig kappa-chain is 400 nM or more at pH8. Pl is  
 CC useful in immunoaffinity chromatography. The present sequence is a  
 CC Peptostreptococcus protein L Ig light chain binding domain, which is  
 CC given in the disclosure of the present invention.  
 XX  
 XX Sequence 72 AA;  
 SQ  
 Query Match 91.6%; Score 339; DB 21; Length 72;  
 Best Local Similarity 90.3%; Pred. No. 5.5e-35;  
 Matches 65; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 KETPEEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKYTDLE 60  
 DB 1 KETPEEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKYTDVA 60  
 QY 61 DGGYTINIRFAG 72  
 DB 61 DKGYTINIRFAG 72  
 RESULT 6  
 AAY82544  
 ID AAY82544 standard; Protein; 75 AA.  
 XX  
 AC AAY82544;  
 XX  
 DT 20-JUL-2000 (first entry)  
 XX  
 DE Peptostreptococcus strain 3316 protein L domain C4 protein sequence.  
 XX  
 KW Immunoglobulin light chain binding protein; PpL; protein L;  
 KW Peptostreptococcus; human immunoglobulin kappa chain;  
 KW immunoaffinity chromatography.  
 XX  
 OS Peptostreptococcus sp.  
 XX  
 PN WO200015803-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 14-SEP-1999; 99WO-GB03048.  
 XX  
 PR 14-SEP-1998; 98GB-0019998.  
 PR 26-APR-1999; 99GB-0009578.  
 XX  
 PA (ACTI-) ACTINOVA LTD.  
 XX  
 XX Gore MG, Beckingham JA, Roberts SE;  
 PI  
 XX WPI; 2000-271441/23.  
 DR N-PSDB; AAA08433.  
 DR  
 XX New modified immunoglobulin light chain binding protein, useful in  
 PT immunoaffinity chromatography, has a dissociation constant of 400 nM or

PT more at pH8 with respect to human immunoglobulin kappa-chain -  
 XX  
 PS Disclosure; Page 49-50; 56pp; English.  
 XX  
 XX The present invention describes an immunoglobulin (Ig) light chain  
 CC binding protein (Pl) which has been modified by one or more amino acid  
 CC substitutions such that the dissociation constant (Kd) of the protein  
 CC with respect to human Ig kappa-chain is 400 nM or more at pH8. Pl is  
 CC useful in immunoaffinity chromatography. The present sequence is a  
 CC Peptostreptococcus protein L Ig light chain binding domain, which is  
 CC given in the disclosure of the present invention.  
 XX  
 XX Sequence 75 AA;  
 SQ  
 Query Match 88.9%; Score 329; DB 21; Length 75;  
 Best Local Similarity 90.0%; Pred. No. 1.1e-33;  
 Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 KTEPEEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKYTDLEDG 62  
 DB 5 ETPEEPKEEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKVNGEYTDLEDG 64  
 QY 63 GYTINIRFAG 72  
 DB 65 GYTINIRFAG 74  
 RESULT 7  
 AAR42204  
 ID AAR42204 standard; Protein; 291 AA.  
 XX  
 AC AAR42204;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 18-MAY-1994 (first entry)  
 XX  
 DE Immunoglobulin binding protein derived from protein L.  
 XX  
 KW Peptide; immunoglobulin; binding; analysis; purification; ELISA;  
 KW enzyme linked immunoabsorbant assay.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9322439-A1.  
 XX  
 PD 11-NOV-1993.  
 XX  
 PF 07-MAY-1993; 93WO-GB00950.  
 XX  
 PR 07-MAY-1992; 92GB-0009804.  
 PR 24-DEC-1992; 92GB-0026928.  
 XX  
 PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.  
 XX  
 XX Atkinson A, Duggieby CJ, Murphy JP, Trowern AR;  
 XX WPI; 1993-368798/46.  
 DR N-PSDB; AAQ50947.  
 XX  
 XX New immunoglobulin binding proteins derived from Protein L -  
 PT which bind immunoglobulin kappa light chains but not albumin or  
 PT cell walls  
 XX  
 PS Claim 12; Figure 2; 28pp; English.  
 XX  
 CC The synthetic immunoglobulin binding proteins derived from protein  
 CC L comprise repeated sequences from protein L which bind  
 CC immunoglobulin kappa light chains. They can be used in protein  
 CC analysis, purification procedures and other biochemical processes e.  
 CC g. ELISA. The synthetic molecules are of particular advantage if  
 CC they are free of regions in protein L which exhibit albumin and cell  
 CC wall binding.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

```

XX SQ Sequence 291 AA;
Query Match 88.9%; Score 329; DB 14; Length 291;
Best Local Similarity 90.0%; Pred. No. 6.5e-33;
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KTPPEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKYTADLEDG 62
Db 503 ETPEEPKEEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKVNGEYTDLEDG 562

Qy 63 GYTINIRFAG 72
Db 563 GYTINIRFAG 572

RESULT 9
AAR43699
ID AAR43699 standard; Protein; 1027 AA.
XX AC AAR43699;
XX DT 25-MAR-2003 (updated)
XX DT 18-MAY-1994 (first entry)
XX DE Protein L.
XX KW Protein; immunoglobulin; binding; immobilisation; light chains;
XX KW antibodies; diagnosis; pharmaceutical; ss.
XX OS Peptococcus magnus.
XX FH Key Location/Qualifiers
XX FT Peptide 36..59
XX FT /label= Signal sequence.
XX FT Protein 60..968
XX FT /label= Mature protein L.
XX PN WO9322438-A1.
XX PD 11-NOV-1993.
XX PF 07-MAY-1993; 93WO-GB00949.
XX PR 07-MAY-1992; 92GB-0009804.
XX PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX PI Atkinson A, Duggleby CJ, Murphy JP, Trowern AR;
XX PR WPI: 1993-368797/46.
XX DR P-PSDB; AAR43699.
XX PA Immunoglobulin binding polypeptide, protein L - used for prodn.
XX PT of pharmaceuticals and for immobilising antibodies e.g. on
XX PT columns, in diagnostic tests and in assays
XX PS Claim 4; Figure 1; 29pp; English.
XX CC Protein L forms a complex with immunoglobulin kappa light chain.
XX CC Purified protein can be used as a reagent for immobilising
XX CC antibodies e.g. on columns, in diagnostic tests and in assays. It
XX CC may also be used in the production of pharmaceuticals.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 1027 AA;
Query Match 88.9%; Score 329; DB 14; Length 1027;
Best Local Similarity 90.0%; Pred. No. 3.6e-32;
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KTPPEPKEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKYTADLEDG 62
Db 503 ETPEEPKEEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKVNGEYTDLEDG 562

Qy 63 GYTINIRFAG 72
Db 563 GYTINIRFAG 572

RESULT 8
AAR42203
ID AAR42203 standard; Protein; 1027 AA.
XX AC AAR42203;
XX DT 25-MAR-2003 (updated)
XX DT 18-MAY-1994 (first entry)
XX DE Protein L.
XX KW Peptide; immunoglobulin; binding; analysis; purification; ELISA;
XX KW enzyme linked immunoabsorbant assay.
XX OS Peptococcus magnus.
XX FH Key Location/Qualifiers
XX FT Peptide 36..59
XX FT /label= Signal sequence.
XX FT Protein 60..968
XX FT /label= Mature protein L.
XX PN WO9322439-A1.
XX PD 11-NOV-1993.
XX PF 07-MAY-1993; 93WO-GB00950.
XX PR 07-MAY-1992; 92GB-0009804.
XX PR 24-DEC-1992; 92GB-0026928.
XX PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX PI Atkinson A, Duggleby CJ, Murphy JP, Trowern AR;
XX PR WPI: 1993-368798/46.
XX DR N-PSDB; AAR42203.
XX PA New immunoglobulin binding proteins derived from Protein L -
XX PT which bind immunoglobulin kappa light chains but not albumin or
XX PT cell walls
XX PS Disclosure; Figure 1; 28pp; English.
XX CC The synthetic immunoglobulin binding proteins derived from protein
XX CC L correspond to the repeated sequences in protein L which bind
XX CC immunoglobulin kappa light chains. They can be used in protein
XX CC analysis, purification procedures and other biochemical processes e.
XX CC g. ELISA. The synthetic molecules are of particular advantage if
XX CC they are free of regions in protein L which exhibit albumin and cell
XX CC wall binding.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 1027 AA;
Query Match 88.9%; Score 329; DB 14; Length 1027;
Best Local Similarity 90.0%; Pred. No. 3.6e-32;

```

```

RESULT 10
AAY82538
ID AAY82538 standard; Protein; 72 AA.
XX
XX AC AAY82538;
XX
DT 20-JUL-2000 (first entry)
XX
XX DE Peptostreptococcus strain 312 protein L domain B2 protein sequence.
XX
XX DE Immunoglobulin light chain binding protein; PpL; protein L;
XX KW Peptostreptococcus; human immunoglobulin kappa chain,
XX KW immunoreactivity chromatography.
XX
XX OS Peptostreptococcus sp.
XX
XX PN WO200015803-A1.
XX
XX PD 23-MAR-2000.
XX
XX PF 14-SEP-1999; 99WO-GB03048.
XX
XX PR 14-SEP-1998; 98GB-0019998.
XX PR 26-APR-1999; 99GB-0009578.
XX
XX PA (ACTI-) ACTINOVA LTD.
XX
XX PI Gore MG, Beckingham JA, Roberts SE;
XX
XX DR WPI; 2000-271441/23.
XX DR N-PSDB; AAA08427.
XX
XX PT New modified immunoglobulin light chain binding protein, useful in
XX PT immunoreactivity chromatography, has a dissociation constant of 400 nM or
XX PT more at pH8 with respect to human immunoglobulin kappa-chain -
XX PS Disclosure; Page 42; 56pp; English.
XX
XX CC The present invention describes an immunoglobulin (Ig) light chain
XX CC binding protein (PI) which has been modified by one or more amino acid
XX CC substitutions such that the dissociation constant (Kd) of the protein
XX CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is
XX CC useful in immunoreactivity chromatography. The present sequence is a
XX CC Peptostreptococcus protein L Ig light chain binding domain, which is
XX CC given in the disclosure of the present invention.
XX
XX SQ Sequence 72 AA;

Query Match 87.0%; Score 322; DB 21; Length 72;
Best Local Similarity 84.7%; Pred. No. 7.6e-33;
Matches 61; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 KETPTPEPREEVITKANLIYADGKTQTAEFGKTFEAETAEAYRYADLLAKENGYTADLE 60
DB 1 KETPTPEPREEVITKANLIYADGKTQTAEFGKTFEAETAEAYRYADALKKNGEYTYDVA 60
QY 61 DGGYTNIRFAG 72
DB 61 DKGYTLNIRFAG 72

RESULT 11
AAB10432
ID AAB10432 standard; Protein; 367 AA.
XX
XX AC AAB10432;
XX
XX DT 01-DEC-2000 (first entry)
XX
XX DE Expression vector pSEX11L4 protein G.
XX

```



XX The present invention describes an immunoglobulin (Ig) light chain  
 CC binding protein (PI) which has been modified by one or more amino acid  
 CC substitutions such that the dissociation constant (Kd) of the protein  
 CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
 CC useful in immunoaffinity chromatography. The present sequence  
 CC represents an Ig light chain binding PpL construct derived from  
 CC Peptostreptococcus sp. protein L.  
 XX  
 SQ Sequence 82 AA;

Query Match 82.2%; Score 304; DB 21; Length 82;  
 Best Local Similarity 84.3%; Pred. No. 1.7e-30;  
 Matches 59; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 KTPPEPKKEVTIKANLIYADGKTQTAEFGKGTAEATAEAYRYADLLAKENGKYTADLEDG 62  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 12 ETPEEPKEEVTIKVNLIFADGKIQTAEFGKGTAEATAEAYRYADLLAKVNGEYTDLEDG 71

QY 63 GYTINIRFAG 72  
 :|||||  
 Db 72 GNHMIKIFAG 81

RESULT 15  
 AAY82545  
 ID AAY82545 standard; Protein; 82 AA.  
 XX  
 AC AAY82545;  
 XX  
 DT 20-JUL-2000 (first entry)  
 XX  
 DE PpL mutant protein sequence SEQ ID NO:19.  
 XX  
 KW Immunoglobulin light chain binding protein; PpL; protein L;  
 KW Peptostreptococcus; human immunoglobulin kappa chain; mutant;  
 KW immunoaffinity chromatography; site directed mutagenesis.  
 XX  
 OS Peptostreptococcus sp.  
 XX  
 PN WO200015803-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 14-SEP-1999; 99WO-GB03048.  
 XX  
 PR 14-SEP-1998; 98GB-0019998.  
 PR 26-APR-1999; 99GB-0009578.  
 XX  
 PA (ACTI-) ACTINOVA LTD.  
 XX  
 PI Gore MG, Beckingham JA, Roberts SE;  
 XX  
 DR WPI; 2000-271441/23.  
 DR N-PSDB; AAA08442.  
 XX  
 PT New modified immunoglobulin light chain binding protein, useful in  
 PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
 PT more at pH8 with respect to human immunoglobulin kappa-chain -  
 XX  
 PS Example 1; Page 50-51; 56pp; English.  
 XX  
 CC The present invention describes an immunoglobulin (Ig) light chain  
 CC binding protein (PI) which has been modified by one or more amino acid  
 CC substitutions such that the dissociation constant (Kd) of the protein  
 CC with respect to human Ig kappa-chain is 400 nM or more at pH8. PI is  
 CC useful in immunoaffinity chromatography. The present sequence is a  
 CC mutant Peptostreptococcus protein L, which is an Ig light chain binding  
 CC protein, from the present invention.  
 XX  
 SQ Sequence 82 AA;

Query Match 80.3%; Score 297; DB 21; Length 82;

Best Local Similarity 82.9%; Pred. No. 1.3e-29;  
 Matches 58; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 KTPPEPKKEVTIKANLIYADGKTQTAEFGKGTAEATAEAYRYADLLAKENGKYTADLEDG 62  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 12 ETPEEPKEEVTIKVNLIFADGKIQTAEFGKGTAEATAEAYRYADLLAKVNGEYTDLEDG 71

QY 63 GYTINIRFAG 72  
 :|||||  
 Db 72 GNHMIKIFAG 81

Search completed: September 3, 2003, 11:19:19  
 Job time : 49.3288 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:16:02 ; Search time 15.7808 Seconds  
(without alignments)  
193.043 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_225\_296

Perfect score: 370

Sequence: 1 KEKTPPEKPEVTIKANLIY.....GKYTADLEGGYTINIRFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370	100.0	305	2	US-08-795-475-1
2	370	100.0	434	2	US-08-795-475-3
3	329	88.9	75	3	US-08-446-137B-8
4	329	88.9	291	3	US-08-446-137B-4
5	329	88.9	1027	3	US-08-446-137B-2
6	315.5	85.3	71	3	US-08-446-137B-6
7	306	82.7	74	3	US-08-446-137B-7
8	254	68.6	182	3	US-08-828-741B-2
9	254	68.6	182	4	US-09-160-567-2
10	254	68.6	182	4	US-09-710-299-2
11	250.5	67.7	71	3	US-08-446-137B-5
12	246	66.5	178	3	US-08-828-741B-13
13	246	66.5	178	4	US-09-160-567-13
14	246	66.5	178	4	US-09-710-299-13
15	246	66.5	198	3	US-08-828-741B-8
16	246	66.5	198	4	US-09-160-567-8
17	246	66.5	198	4	US-09-710-299-8
18	246	66.5	342	3	US-08-828-741B-6
19	246	66.5	342	4	US-09-160-567-6
20	246	66.5	342	4	US-09-710-299-6
21	246	66.5	495	3	US-08-828-741B-4
22	246	66.5	495	4	US-09-160-567-4
23	246	66.5	495	4	US-09-710-299-4
24	65	17.6	342	2	US-08-724-394A-6
25	65	17.6	540	2	US-08-724-394A-4
26	65	17.6	610	2	US-08-724-394A-5
27	64.5	17.4	279	4	US-09-314-701-60

28	62	16.8	490	4	US-09-252-149B-26	Sequence 26, Appl
29	62	16.8	544	1	US-08-387-156-10	Sequence 10, Appl
30	62	16.8	544	2	US-08-694-865-10	Sequence 10, Appl
31	62	16.8	544	2	US-08-878-748-10	Sequence 10, Appl
32	62	16.8	544	3	US-09-124-491-10	Sequence 10, Appl
33	62	16.8	544	4	US-09-383-912-10	Sequence 10, Appl
34	62	16.8	699	2	US-08-694-865-16	Sequence 16, Appl
35	62	16.8	699	3	US-09-124-491-16	Sequence 16, Appl
36	62	16.8	699	4	US-09-383-912-16	Sequence 16, Appl
37	62	16.8	770	4	US-09-252-991A-26865	Sequence 26865, A
38	62	16.8	924	3	US-08-619-812-8	Sequence 8, Appl
39	62	16.8	926	1	US-07-908-253-2	Sequence 2, Appl
40	62	16.8	926	1	US-08-455-970A-2	Sequence 2, Appl
41	62	16.8	926	1	US-08-387-156-6	Sequence 6, Appl
42	62	16.8	926	2	US-08-694-865-6	Sequence 6, Appl
43	62	16.8	926	2	US-08-878-748-6	Sequence 6, Appl
44	62	16.8	926	2	US-08-535-837-2	Sequence 2, Appl
45	62	16.8	926	3	US-09-124-491-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-08-795-475-1  
; Sequence 1, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bivrock, Lars  
; APPLICANT: Sjvdrick, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
US-08-795-475-1

Query Match 100.0%; Score 370; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 5,le-40;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEKPEVTIKANLIYADGKQTQAEKGTFAEATYADLAKENGYTABLE 60  
|||||

Db 225 KEKTPPEKVEVTKANLIYADGKTQTAEKGTFAEATAEAYRYADLLAKENGYTADLE 284  
 Qy 61 DGGYTINIRFAG 72  
 Db 285 DGGYTINIRFAG 296

## RESULT 2

US-08-795-475-3  
 ; Sequence 3, Application US/08795475  
 ; Patent No. 5965390  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bjvick, Lars  
 ; APPLICANT: Sjdbring, Ulf  
 ; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104-7092  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/795,475  
 ; FILING DATE: 11-FEB-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McMasters, David D.  
 ; REGISTRATION NUMBER: 33,963  
 ; REFERENCE/DOCKET NUMBER: 100084.402D1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 434 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: NO  
 ; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
 ; US-08-795-475-3

Query Match 100.0%; Score 370; DB 2; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-40;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KEKTPPEKVEVTKANLIYADGKTQTAEKGTFAEATAEAYRYADLLAKENGYTADLE 60  
 Db 225 KEKTPPEKVEVTKANLIYADGKTQTAEKGTFAEATAEAYRYADLLAKENGYTADLE 284  
 Qy 61 DGGYTINIRFAG 72  
 Db 285 DGGYTINIRFAG 296

## RESULT 3

US-08-446-137B-8  
 ; Sequence 8, Application US/08446137B  
 ; Patent No. 6162903  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Towerin, Angus R.  
 ; APPLICANT: Atkinson, Anthony  
 ; APPLICANT: Murphy, Jonathan P.  
 ; APPLICANT: Laurence, Oliver S.

; APPLICANT: Duggleby, Clive J.  
 ; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
 ; FROM L PROTEIN AND THEIR USES  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/446,137B  
 ; FILING DATE: 22-MAY-1995  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McMasters, David D.  
 ; REGISTRATION NUMBER: 33,963  
 ; REFERENCE/DOCKET NUMBER: 100084.406  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 75 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; US-08-446-137B-8

Query Match 88.9%; Score 329; DB 3; Length 75;  
 Best Local Similarity 90.0%; Pred. No. 1.6e-35;  
 Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Qy 3 KTEPEKVEVTKANLIYADGKTQTAEKGTFAEATAEAYRYADLLAKENGYTADLE 62  
 Db 5 ETPPEKVEVTKVNLIFADGKTQTAEKGTFAEATAEAYRYADLLAKENGYTADLE 64  
 Qy 63 GYTINIRFAG 72  
 Db 65 GYTINIRFAG 74

RESULT 4  
 US-08-446-137B-4  
 ; Sequence 4, Application US/08446137B  
 ; Patent No. 6162903  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Towerin, Angus R.  
 ; APPLICANT: Atkinson, Anthony  
 ; APPLICANT: Murphy, Jonathan P.  
 ; APPLICANT: Laurence, Oliver S.  
 ; APPLICANT: Duggleby, Clive J.  
 ; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
 ; FROM L PROTEIN AND THEIR USES  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-137B-4

Query Match      88.9%; Score 329; DB 3; Length 291;
Best Local Similarity 90.0%; Pred. No. 1e-34;
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      3 KTEPEPKEEVIKANLIYADGKTQTAEFGKTFEAATAEAYRYADLLAKENGYTADLEDG 62
Db      221 ETPEPKEEVIKVNLIADFADGKTQTAEFGKTFEAATAEAYRYADLLAKVNGEYTDLEDG 280

QY      63 GYTINIRFAG 72
Db      281 GYTINIRFAG 290

RESULT 5
US-08-446-137B-2
; Sequence 2, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; TITLE OF INVENTION: FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-446-137B-2

Query Match      88.9%; Score 329; DB 3; Length 1027;
Best Local Similarity 90.0%; Pred. No. 5.7e-34;
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      3 KTEPEPKEEVIKANLIYADGKTQTAEFGKTFEAATAEAYRYADLLAKENGYTADLEDG 62
Db      503 ETPEPKEEVIKVNLIADFADGKTQTAEFGKTFEAATAEAYRYADLLAKVNGEYTDLEDG 562

QY      63 GYTINIRFAG 72
Db      563 GYTINIRFAG 572

RESULT 6
US-08-446-137B-6
; Sequence 6, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; TITLE OF INVENTION: FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-446-137B-6

Query Match      85.3%; Score 315.5; DB 3; Length 71;
Best Local Similarity 87.5%; Pred. No. 8.5e-34;
Matches 63; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY      1 KEKTEPEPKEEVIKANLIYADGKTQTAEFGKTFEAATAEAYRYADLLAKENGYTADLE 60
Db      1 KEK-PEEPKEEVIKVNLIADFADGKTQTAEFGKTFEAATAEAYRYADLLAKENGYTADLE 59

QY      61 DGGYTINIRFAG 72
Db      60 DGGYTINIRFAG 71
```

RESULT 7  
 US-08-446-137B-7  
 ; Sequence 7, Application US/08446137B  
 ; Patent No. 6162903  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tower, Angus R.  
 ; APPLICANT: Atkinson, Anthony  
 ; APPLICANT: Murphy, Jonathan P.  
 ; APPLICANT: Laurence, Oliver S.  
 ; APPLICANT: Dugleby, Clive J.  
 ; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
 ; FROM L PROTEIN AND THEIR USES  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED AND BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/446,137B  
 ; FILING DATE: 22-MAY-1995  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mcmasters, David D.  
 ; REGISTRATION NUMBER: 33,963  
 ; REFERENCE/DOCKET NUMBER: 100084.406  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 74 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ;  
 ; US-08-446-137B-7  
 ;  
 Query Match 82.7%; Score 306; DB 3; Length 74;  
 Best Local Similarity 84.3%; Pred. No. 1.5e-32;  
 Matches 59; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 KTPPEKVEVTIKANLIYADGKTQTAEFGTFAETAAYRYADLLAKENGKYTDLEDG 62  
 Db 5 ETPPEKEEVIKNLFIADGKIQTAEFGTFAETAAYRYADLLAKENGKYTDLEDG 64  
 QY 63 GTYTINIRFAG 72  
 Db 65 GNTINIKFAG 74  
 RESULT 8  
 US-08-828-741B-2  
 ; Sequence 2, Application US/08828741B  
 ; Patent No. 6043069  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koentgen, Frank  
 ; APPLICANT: Suess, Gabriele M.  
 ; APPLICANT: Tarlinton, David M.  
 ; APPLICANT: Treutlein, Herbert R.  
 ; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
 ; PRODUCING SAME  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
 ; STREET: 400 Garden City Plaza  
 ; CITY: Garden City

[illegible]

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-710-299-2

Query Match      68.6%; Score 254; DB 4; Length 182;
Best Local Similarity 66.2%; Pred. No. 3e-25;
Matches 47; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY  2  EKTPEPKKEVTKANLIYADGKTQTAERFKGTFAETAAEAYRYADLLAKENGKYTADLEDG 61
Db   20  QAAQPKNTEVTKANLIFANGSTQAEFKGTTEKATSEAYAYADTLKDNGETYVDVAD 79
QY   62  GGYTINIRFAG 72
Db   80  KGITLNIKFAG 90
      |||:|||||
      |||:|||||

RESULT 11
US-08-446-137B-5
; Sequence 5, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Dugleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-446-137B-5

Query Match      67.7%; Score 250.5; DB 3; Length 71;
Best Local Similarity 66.6%; Pred. No. 2.4e-25;
Matches 48; Conservative 11; Mismatches 10; Indels 1; Gaps 1;

QY  3  KTPPEKEEYTIKANLIYADGKTQTAERFKGTFAETAAEAYRYADLLAKENGKYTADLEDG 62
Db   2  ETP-EPEEYTIKANLIFADGSTQAEFKGTTEKATSEAYAYADTLKDNGETYVDVADK 60
QY   63  GYTINIRFAG 72
Db   61  GLTINIKFAG 70
      |:|:|:|:|
      |:|:|:|:|

```

RESULT 12  
US-08-828-741B-13  
; Sequence 13, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,741B  
; FILING DATE: 26-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 178 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-828-741B-13

Query Match 66.5%; Score 246; DB 3; Length 178;  
Best Local Similarity 73.0%; Pred. No. 3.2e-24;  
Matches 46; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 10 EEVTKANLIYDGKQTQAEFGKTFAEATAYRYADLLAKENGKYTADLEDGGYTINIR 69  
Db 12 EEVTKANLIIFANGSTQAEFGKTFEKATSEAYAYADTLKKNGEYTVDVADKGYTLNIK 71  
QY 70 FAG 72  
Db 72 FAG 74

RESULT 13  
US-09-160-567-13  
; Sequence 13, Application US/09160567  
; Patent No. 6326179  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City

STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-13  
Query Match 66.5%; Score 246; DB 4; Length 178;  
Best Local Similarity 73.0%; Pred. No. 3.2e-24;  
Matches 46; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 10 EEVTKANLIYDGKQTQAEFGKTFAEATAYRYADLLAKENGKYTADLEDGGYTINIR 69  
Db 12 EEVTKANLIIFANGSTQAEFGKTFEKATSEAYAYADTLKKNGEYTVDVADKGYTLNIK 71  
QY 70 FAG 72  
Db 72 FAG 74  
RESULT 14  
US-09-710-299-13  
; Sequence 13, Application US/09710299  
; Patent No. 6521741  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/710,299  
; FILING DATE: 09-NOV-6521741-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: 08/828,741
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-710-299-13

Query Match          66.5%; Score 246; DB 4; Length 178;
Best Local Similarity 73.0%; Pred. No. 3.2e-24;
Matches 46; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 10 EETIKANLIYADGKTQTAEFKGTFAETAETAEAYRYADLLAKENGYTADLEGGYTINIR 69
Db 12 EETIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKONGEYTVDVADKGYTLNIK 71
QY 70 FAG 72
Db 72 FAG 74

RESULT 15
US-08-828-741B-8
; Sequence 8, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; NUMBER OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
US-08-828-741B-8

Query Match          66.5%; Score 246; DB 3; Length 198;
Best Local Similarity 73.0%; Pred. No. 3.7e-24;
Matches 46; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 10 EETIKANLIYADGKTQTAEFKGTFAETAETAEAYRYADLLAKENGYTADLEGGYTINIR 69
Db 32 EETIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKONGEYTVDVADKGYTLNIK 91
QY 70 FAG 72
Db 92 FAG 94

Search completed: September 3, 2003, 11:25:37
Job time : 16.7808 secs
```

THIS PAGE BLANK (USPTO)





```
; APPLICANT: Bjvrck, Lars
; APPLICANT: Sjvbrng, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; US-08-325-278-1

Query Match 100.0%; Score 370; DB 8; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.2e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTDLE 60
Db 225 KKTPEEPKEEVTKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTDLE 60

QY 61 DGGYTINIRFAG 72
Db 285 DGGYTINIRFAG 296

RESULT 3
US-08-325-278-3
; Sequence 3, Application US/08325278
; Publication No. US20030027283A1
; GENERAL INFORMATION:
; APPLICANT: Bjvrck, Lars
; APPLICANT: Sjvbrng, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055
; US-08-325-278-3

Query Match 100.0%; Score 370; DB 8; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.8e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTDLE 60
Db 225 KKTPEEPKEEVTKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTDLE 60

QY 61 DGGYTINIRFAG 72
Db 285 DGGYTINIRFAG 296

RESULT 4
US-08-808-212A-8
; Sequence 8, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Stan Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
; US-08-808-212A-8

Query Match 91.6%; Score 339; DB 10; Length 72;
Best Local Similarity 90.3%; Pred. No. 8.4e-33;
Matches 65; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTDLE 60
Db 1 KKTPEEPKEEVTKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTDVA 60

QY 61 DGGYTINIRFAG 72
Db 61 DGGYTINIRFAG 72

RESULT 5
US-08-808-212A-18
; Sequence 18, Application US/09808212A
; Patent No. US20020137918A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 18
; TYPE: PRT
; LENGTH: 75
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-18

Query Match      88.9%; Score 329; DB 10; Length 75;
Best Local Similarity 90.0%; Pred. No. 1.3e-31;
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KTEPEEKKEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGYTADLEDG 62
Db 5 ETPEEKKEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKVNGEYTDLEDG 64

Qy 63 GYTINIRFAG 72
Db 65 GYTINIRFAG 74

RESULT 6
US-09-808-212A-6
; Sequence 6, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-6

Query Match      87.0%; Score 322; DB 10; Length 72;
Best Local Similarity 84.7%; Pred. No. 8.5e-31;
Matches 61; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KTEPEEKKEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGYTADLE 60
Db 1 KTEPEEKKEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADALKDNGEYTDVA 60

Qy 61 DGGYTINIRFAG 72
Db 61 DGGYTINIRFAG 72

RESULT 7
US-09-808-212A-14
; Sequence 14, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
```

```
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-14

Query Match      85.3%; Score 315.5; DB 10; Length 71;
Best Local Similarity 87.5%; Pred. No. 4.8e-30;
Matches 63; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 1 KTEPEEKKEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGYTADLE 60
Db 1 KEK-PEEPKEEVIKVNLIADGKTQTAEFGKTFEATAKAYAYADLLAKENGEYTDLE 59

Qy 61 DGGYTINIRFAG 72
Db 60 DGGTINIRFAG 71

RESULT 8
US-09-808-212A-16
; Sequence 16, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-16

Query Match      82.7%; Score 306; DB 10; Length 74;
Best Local Similarity 84.3%; Pred. No. 6.7e-29;
Matches 59; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 3 KTEPEEKKEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGYTADLEDG 62
Db 5 ETPEEKKEVTIKVNLIFADGKTQTAEFGKTFEATAKAYAYANLLAKENGEYTDLEDG 64

Qy 63 GYTINIRFAG 72
Db 65 GNTINIRFAG 74

RESULT 9
US-09-808-212A-2
; Sequence 2, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
```

## US-09-808-212A-2

Query Match 82.2%; Score 304; DB 10; Length 82;  
Best Local Similarity 84.3%; Pred. No. 1.3e-28;  
Matches 59; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 3 KTEPEKPEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTADLEDG 62  
DB 12 ETPEPEKPEVTIKVNLIFADGKIQTAEFGTFAEATAEAYRYADLLAKVNGEYTDLEDG 71  
QY 63 GYTINIRFAG 72  
DB 72 GNMNIKFAG 81

## RESULT 10

US-09-808-212A-4  
; Sequence 4, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-4

Query Match 72.7%; Score 269; DB 10; Length 76;  
Best Local Similarity 68.4%; Pred. No. 1.6e-24;  
Matches 52; Conservative 11; Mismatches 9; Indels 4; Gaps 1;  
QY 1 KKTPEP-----KEEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYT 56  
DB 1 KEETPETDSEEEVTIKANLIFANGSTQAEFGTFAEATAEAYRYADLLAKNGEY 60  
QY 57 ADLEGGYTINIRFAG 72  
DB 61 VDVAADGGYTINIRFAG 76

## RESULT 11

US-10-345-618-2  
; Sequence 2, Application US/10345618  
; Publication No. US2003014848A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:LHL protein  
; OTHER INFORMATION: sequence  
US-10-345-618-2

Query Match 68.6%; Score 254; DB 12; Length 182;  
Best Local Similarity 66.2%; Pred. No. 2.9e-22;  
Matches 47; Conservative 12; Mismatches 12; Indels 0; Gaps 0;  
QY 2 EKTPEPEKPEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTADLED 61  
DB 20 QAAPKNDTEEVTKANLIFANGSTQAEFGTFAEATAEAYRYADLLAKNGEYTDVAD 79  
QY 62 GYTINIRFAG 72  
DB 80 KGYTLNIRFAG 90

## RESULT 12

US-09-808-212A-12  
; Sequence 12, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 71  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-12

Query Match 68.0%; Score 251.5; DB 10; Length 71;  
Best Local Similarity 67.6%; Pred. No. 1.7e-22;  
Matches 48; Conservative 12; Mismatches 10; Indels 1; Gaps 1;  
QY 2 EKTPEPEKPEVTIKANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGKYTADLED 61  
DB 1 KETP-EPEEVTIKANLIFADGSTQAEFGTFAEATAEAYRYADLLAKNGEYTDVAD 59  
QY 62 GYTINIRFAG 72  
DB 60 KGLTLNIRFAG 70

## RESULT 13

US-10-345-618-13  
; Sequence 13, Application US/10345618  
; Publication No. US2003014848A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:LHL-omp  
; OTHER INFORMATION: protein sequence  
US-10-345-618-13





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:59:36 ; Search time 16.4438 Seconds  
(without alignments)  
1116.704 Million cell updates/sec

Title: US-08-325-278B-3

Perfect score: 2235

Sequence: 1 AVENKETPTPTDSEEV.....GVDGWTVDATKFTVTM 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	100.0	434	2	US-08-795-475-3
2	1565	70.0	305	2	US-08-795-475-1
3	1263.5	56.5	1027	3	US-08-446-137B-2
4	1216	54.4	291	3	US-08-446-137B-4
5	706.5	31.6	664	3	US-08-669-408B-2
6	622.5	27.9	502	1	US-08-378-761A-25
7	622.5	27.9	502	1	US-08-485-286-25
8	622	27.8	493	1	US-08-378-761A-23
9	622	27.8	493	1	US-08-485-286-23
10	618	27.7	342	3	US-08-828-741B-6
11	618	27.7	342	4	US-09-160-567-6
12	618	27.7	342	4	US-09-710-299-6
13	618	27.7	489	1	US-08-378-761A-19
14	618	27.7	489	1	US-08-485-286-19
15	615	27.5	178	3	US-08-828-741B-13
16	615	27.5	178	4	US-09-160-567-13
17	615	27.5	178	4	US-09-710-299-13
18	615	27.5	198	3	US-08-828-741B-8
19	615	27.5	198	4	US-09-160-567-8
20	615	27.5	198	4	US-09-710-299-8
21	614	27.5	495	3	US-08-828-741B-4
22	614	27.5	495	4	US-09-160-567-4
23	614	27.5	495	4	US-09-710-299-4
24	613	27.4	182	3	US-08-828-741B-2
25	613	27.4	182	4	US-09-160-567-2
26	613	27.4	182	4	US-09-710-299-2
27	404	18.1	413	3	US-08-669-408B-10

28	369.5	16.5	402	1	US-08-378-761A-17	Sequence 17, Appl
29	369.5	16.5	402	1	US-08-485-286-17	Sequence 8, Appl
30	334	14.9	75	3	US-08-446-137B-8	Sequence 6, Appl
31	315.5	14.1	71	3	US-08-446-137B-6	Sequence 7, Appl
32	306	13.7	74	3	US-08-446-137B-7	Sequence 5, Appl
33	303	13.6	71	3	US-08-446-137B-5	Sequence 2, Appl
34	293	13.1	57	3	US-09-117-233-16	Sequence 12, Appl
35	292	13.1	57	3	US-09-117-233-2	Sequence 38, Appl
36	292	13.1	60	3	US-09-117-233-12	Sequence 10, Appl
37	287	12.8	56	3	US-09-058-459-38	Sequence 6, Appl
38	287	12.8	56	3	US-09-127-926-38	Sequence 4, Appl
39	286	12.8	57	3	US-09-117-233-8	Sequence 68, Appl
40	286	12.8	58	3	US-09-117-233-10	Sequence 58, Appl
41	286	12.8	58	3	US-09-444-305A-1	
42	286	12.8	60	3	US-09-117-233-6	
43	285	12.8	60	3	US-09-117-233-4	
44	283	12.7	56	3	US-09-058-459-68	
45	283	12.7	56	3	US-09-127-926-68	

ALIGNMENTS

RESULT 1  
US-08-795-475-3  
; Sequence 3, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bjvick, Lars  
; APPLICANT: Sjvbring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
US-08-795-475-3

Query Match 100.0%; Score 2235; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 6.2e-170;  
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKETPTPTDSEEVITKANLIFANGSTQAEKTFEKAISEAYADTLKKDN 60  
|||||

Db 1 AVENKEETPETDSEEEVITKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKN 60  
QY 61 GETVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTAFFKGTPEATAE 120  
Db 61 GETVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTAFFKGTPEATAE 120  
QY 121 AVYADALKDNGEYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTA 180  
Db 121 AVYADALKDNGEYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTA 180  
QY 181 EFKGTFEEATAEAYRYADLLAKENGYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKA 240  
Db 181 EFKGTFEEATAEAYRYADLLAKENGYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKA 240  
QY 241 NLIYADGKTQTAFFKGTPEATAEAYRYADLLAKENGYTADLEGGYTNIRFAGKVD 300  
Db 241 NLIYADGKTQTAFFKGTPEATAEAYRYADLLAKENGYTADLEGGYTNIRFAGKVD 300  
QY 301 EKPEEPMDTYKLLNGKTLKGETTTEAVDAATAEKVFKOYANDNGVDGWTYDDATKTF 360  
Db 301 EKPEEPMDTYKLLNGKTLKGETTTEAVDAATAEKVFKOYANDNGVDGWTYDDATKTF 360  
QY 361 VTEKPEVIDASELTAVTYTKLVINGKTLKGETTKAVDAETAERAFKQYANDNGVDGYW 420  
Db 361 VTEKPEVIDASELTAVTYTKLVINGKTLKGETTKAVDAETAERAFKQYANDNGVDGYW 420  
QY 421 TYDDATKTFVTTEM 434  
Db 421 TYDDATKTFVTTEM 434

## RESULT 2

US-08-795-475-1  
; Sequence 1, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bjvrck, Lars  
; APPLICANT: Bjvrck, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

## US-08-795-475-1

Query Match 70.0%; Score 1565; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 6.9e-117;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AVENKEETPETDSEEEVITKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKN 60  
Db 1 AVENKEETPETDSEEEVITKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKN 60  
QY 61 GETVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTAFFKGTPEATAE 120  
Db 61 GETVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTAFFKGTPEATAE 120  
QY 121 AVYADALKDNGEYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTA 180  
Db 121 AVYADALKDNGEYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKANLIYADGKTQTA 180  
QY 181 EFKGTFEEATAEAYRYADLLAKENGYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKA 240  
Db 181 EFKGTFEEATAEAYRYADLLAKENGYTVDVADKGYTLNIFKAGKEKTPPEPKEEVTIKA 240  
QY 241 NLIYADGKTQTAFFKGTPEATAEAYRYADLLAKENGYTADLEGGYTNIRFAGKVD 300  
Db 241 NLIYADGKTQTAFFKGTPEATAEAYRYADLLAKENGYTADLEGGYTNIRFAGKVD 300  
QY 301 EKPEE 305  
Db 301 EKPEE 305

## RESULT 3

US-08-446-137B-2  
; Sequence 2, Application US/08446137B  
; Patent No. 6162903  
; GENERAL INFORMATION:  
; APPLICANT: Townen, Angus R.  
; APPLICANT: Atkinson, Anthony  
; APPLICANT: Murphy, Jonathan P.  
; APPLICANT: Laurence, Oliver S.  
; APPLICANT: Duggleby, Clive J.  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; FROM L PROTEIN AND THEIR USES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,137B  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.406  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1027 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein





```

; REFERENCE/DOCKET NUMBER: 61743/102
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-669-408B-2
;
;
; Query Match 31.6%; Score 706.5; DB 3; Length 664;
; Best Local Similarity 43.9%; Pred. No. 3.6e-48;
; Matches 192; Conservative 51; Mismatches 143; Indels 51; Gaps 17;
;
; QY 24 ANLIFANGSTQTAAEKFGTFPEKATSEAYAYADTLK-KDNGEYTVDVADKGYTL--NIKFA- 79
; DB 148 ANEIVNNSDAYTAESIOPYKLINDAY---DVLSEKDYSKY--SDQKVNVLADQLRDV 202
;
; QY 80 -----GKEKTPPEEKPEEVTIRANLIYADGKTQTAEFGTFPEATAEA--YRYA 125
; DB 203 QAVQLEAPTVIDAPELTPALTYTKLVKVGNTF--SGETTTK-----AIDTATAEKEFKQYA 256
;
; QY 126 DALKKDNGEYTVDVADKGYTLNKA--GKEKTPPEEKPEEVTIRANLIYADGKTQTAEF 182
; DB 257 TANVD--GWSYDDATKTFTVTEKPAVIDAPELTPALTYTKLVKGNTP--SGETTT--- 310
;
; QY 183 KGTFPEATAEAYRYADLLAKENGYTVDVADKGYTLNKA---GKEKTPPEEKPEEVTIK 239
; DB 311 KAVDAETAFAKFAQYATANNVDGWSYDDATKTFTVTEKPAVIDAPELTPALTYTKLVK 370
;
; QY 240 ANLIYADGKTQTAEKFGTFPEATAEAYRYADLLAKENGYTADLEDGGYTNIRFAGK 299
; DB 371 GNTP--SGETTTKAIADAETAKEFKQYATANG---DGWSYDDATKTFTVTEKPA---V 422
;
; QY 300 DERPE--EPMDTYTKILNGKTLKGETTEAVDADAATAEKFQYQYANDNGVDGEWYDDATK 357
; DB 423 IDAPELTPALTYTKLVKNGTFSGETTTKAVDAETAFAKFAQYANENGWYGEWSYDDATK 482
;
; QY 358 TFTVTEKPEVIDASELTPAVTYTKLVINGKTLKGETTTKAVDAETAFAKFAQYANDNGVD 417
; DB 483 TFTVTEKPAVIDAPELTPALTYTKLVINGKTLKGETTTKAVDAETAFAKFAQYANENGVD 542
;
; QY 418 GWYTYDDATKTFTVTM 434
; DB 543 GWYTYDDATKTFTVTM 559
;
;
; RESULT 6
; US-08-378-761A-25
; Sequence 25, Application US/08378761A
; Patent No. 5635384
;
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,286  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/378761  
FILING DATE: 26-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-25

Query Match 27.9%; Score 622.5; DB 1; Length 502;  
Best Local Similarity 41.1%; Pred. No. 1.2e-41;  
Matches 178; Conservative 36; Mismatches 136; Indels 83; Gaps 16;

QY 37 EFKGTF-----EKATSEAYAYADTLKKNNGEYTVDV-ADKGYTLNIFAGKEKTPPEPK 89  
DB 38 DHKGFQPVLPPEKKVPPELWFYTE-LKTRTSITLAIRMDNLYLVGFRTPG----- 87  
QY 90 EEVTKANLIYADGKTQTAEFKGTFEETAEAYRYADALKKNGEYTVDVADKGYTLNLIK 149  
DB 88 -----GYWBEFGKDGDTLLGDNPRWLFGGGRYQD-LIGNKGLTVMGRAEMTRAVN 139  
QY 150 FAGKEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADALKKNGEYTV 209  
DB 140 DLAKKKKAADPO-----ADTKSKLVK-----LVVMVCEGLRFTV-----SRTV 178  
QY 210 DV---ADKGYTLNIFAGKEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTFAEATAEA 265  
DB 179 DAGFNSQHGVTLTVT-OGK---QVKWDRISKAFAEWADHPTAVIPDMQKLGKDKNEA 233  
QY 266 YRYADLLAKENKGYTADLEDGGYTNIRFAGKKVDEKPE-----EPMDTYKLLNKGK 317  
DB 234 ARIVALV--KNQTTAAATAGSCARVRSSCGVD-KPEVIDASELTPAVTTYKLIVNGK 290  
QY 318 TLKGETTTEAVDAATAEKVKFYANDNGVDGWTYDDATKTTTVE----- 363  
DB 291 TLKGETTTEAVDAATAEKVKFYANDNGVDGWTYDDATKTTTVEKPEVIDASELTPAV 350  
QY 364 ---KPEVIDASELTPAVTTYKLIVNGKTLKGETTTKAVDAETAERAKFYANDNGVDGVW 420  
DB 351 TRSKPEVIDASELTPAVTTYKLIVNGKTLKGETTTEAVDAATAEKVKFYANDNGVDGW 410  
QY 421 TYDDATKTTTVE 433  
DB 411 TYDDATKTTTVE 423

## RESULT 8

US-08-378-761A-23  
Sequence 23, Application US/08378761A  
Patent No. 5635384  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
TITLE OF INVENTION: USING  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 493 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-23

Query Match 27.8%; Score 622; DB 1; Length 493;  
Best Local Similarity 40.6%; Pred. No. 1.2e-41;  
Matches 176; Conservative 38; Mismatches 127; Indels 92; Gaps 17;

QY 37 EFKGTF-----EKATSEAYAYADTLKKNNGEYTVDV-ADKGYTLNIFAGKEKTPPEPK 89  
DB 38 DHKGFQPVLPPEKKVPPELWFYTE-LKTRTSITLAIRMDNLYLVGFRTPG----- 87  
QY 90 EEVTKANLIYADGKTQTAEFKGTFEETAEAYRYADALKKNGEYTVDVADKGYTLNLIK 149  
DB 88 -----GYWBEFGKDGDTLLGDNPRWLFGGGRYQD-LIGNKGLTVMGRAEMTRAVN 139  
QY 150 FAGKEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADALKKNGEYTV 209  
DB 140 DLAKKKKAADPO-----ADTKSKLVK-----LVVMVCEGLRFTV-----SRTV 178  
QY 210 DV---ADKGYTLNIFAGKEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTFAEATAEA 265  
DB 179 DAGFNSQHGVTLTVT-OGK---QVKWDRISKAFAEWADHPTAVIPDMQKLGKDKNEA 233  
QY 266 YRYADLLAKENKGYTADLEDGGYTNIRFAGKKVDEKPE-----EPMDTYKLLNKGK 317  
DB 234 ARIVALVKNQ---TAAATAG-SVNV-----DKPEVIDASELTPAVTTYKLIVNGK 281  
QY 318 TLKGETTTEAVDAATAEKVKFYANDNGVDGWTYDDATKTTTVE----- 363  
DB 282 TLKGETTTEAVDAATAEKVKFYANDNGVDGWTYDDATKTTTVEKPEVIDASELTPAV 341  
QY 364 ---KPEVIDASELTPAVTTYKLIVNGKTLKGETTTKAVDAETAERAKFYANDNGVDGVW 420  
DB 342 TRSKPEVIDASELTPAVTTYKLIVNGKTLKGETTTEAVDAATAEKVKFYANDNGVDGW 401  
QY 421 TYDDATKTTTVE 433  
DB 402 TYDDATKTTTVE 414

## RESULT 9

US-08-485-286-23  
; Sequence 23, Application US/08485286  
; Patent No. 5646026  
; Patent No. 5646026 5646119  
; GENERAL INFORMATION:  
; APPLICANT: WALSH, TERENCE A  
; APPLICANT: HEY, TIMOTHY D  
; APPLICANT: MORGAN, ALICE ER  
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
; PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
; TITLE OF INVENTION: USING  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ANDREA T. BORUCKI  
; STREET: 9330 ZIONSVILLE ROAD  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: US  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,286  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/378761  
; FILING DATE: 26-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BORUCKI, ANDREA T  
; REGISTRATION NUMBER: 33651  
; REFERENCE/DOCKET NUMBER: 38272B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 337-4846  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 493 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-485-286-23  
Query Match 27.8%; Score 622; DB 1; Length 493;  
Best Local Similarity 40.6%; Pred. NO. 1.2e-41;  
Matches 176; Conservative 38; Mismatches 127; Indels 92; Gaps 17;  
QY 37 EFKGTF-----ERATSEAYADTLKKDNGEYTDV-ADKGYTLNKFAGKEKTPPEPK 89  
DB 38 DHKGFQVLPPEKVPPELWYTE-LKTRTSITLAIRMDNLYLVGFRTPG-----87  
QY 90 BEVTIKANLIYADGKTQTAEFGTFEEATAEAYRYADALKKDNGETYTDVADKGYTLN 149  
DB 88 -----GVWFEGKDGTHLLGDNPNRWLGFGGRYQD-LIGNKGLTYTMGRAETRAVN 139  
QY 150 FAGKEKTPPEKKEVTIKANLIYADGKTQTAEFGTFEEATAEAYRYADLLAKENGKTV 209  
DB 140 DLAKKKAADPQ-----ADTKSLVK-----LVVMVCEGLRFNTV-----SRTV 178  
QY 210 DV---ADKGYTLNKFAGKEKTPPEKKEVTIKANLIYADGKTQT-AEFKTFEATAEA 265  
DB 179 DAGNSQHGVTLTVT-QGK-----QVQKWDRIKAAFEWADHPTAIPDMOKLGDKNREA 233  
QY 266 YRYADLLAKENGKTYTADLDGGYTLNIRFAGKVDKEPE-----EPMDTYKLILNGK 317  
DB 234 ARIVALKVNTQ---TAAATAG-SVNV-----DKPEVIDASELTPAVTYKLIVNGK 281  
QY 318 TLKGETTTEAADAETAEKVFQYANDNGVDGWTYDDATKTFVTVE-----363  
DB 282 TLKGETTTEAADAETAEKVFQYANDNGVDGWTYDDATKTFVTVEKPEVIDASELTPAV 341

QY 364 ---KPEVIDASELTPAVTYKLIVNGKTLKGETTTKAVDAETAEKAFKOYANDNGVDGW 420  
DB 342 TRSKPEVIDASELTPAVTYKLIVNGKTLKGETTTEAVDAATAEKVFQYANDNGVDGEW 401  
QY 421 TYDDATKTFVTVE 433  
DB 402 TYDDATKTFVTVE 414  
RESULT 10  
US-08-828-741B-6  
; Sequence 6, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,741B  
; FILING DATE: 26-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 342 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-828-741B-6  
Query Match 27.7%; Score 618; DB 3; Length 342;  
Best Local Similarity 75.9%; Pred. NO. 1.5e-41;  
Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;  
QY 16 SEEVITKANLIFANGSTQTAEFGTFEAKTSEAYAYADTLKKONGEYTVDVADKGYTLN 75  
DB 174 SAEVITKANLIFANGSTQTAEFGTFEAKTSEAYAYADTLKKONGEYTVDVADKGYTLN 233  
QY 76 IKFAGKEKTPPE-----PKEVITKANLIYADGKTQTAEFGTFE 115  
DB 234 IKFAGKEATNRTDGTGYILQINSRMGGLTSAEVITKANLIFANGSTQTAEFGTFE 293  
QY 116 EATAEAYRYADALKKDNGETYTDVADKGYTLNIRFAGKEKTPPEPK 161  
DB 294 KATSEAYAYADTLKKDNGETYTDVADKGYTLNIRFAGKESAWRHQ 339

RESULT 11  
US-09-160-567-6  
; Sequence 6, Application US/09160567

Patent No. 6326179  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
APPLICANT: Suess, Gabriele M.  
APPLICANT: Tarlinton, David M.  
APPLICANT: Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
TITLE OF INVENTION: PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-6

Query Match 27.7%; Score 618; DB 4; Length 342;  
Best Local Similarity 75.9%; Pred. No. 1.5e-41;  
Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;

QY 16 SEEEVTKANLIFANGSTQAEFGTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 75  
Db 174 SAEVTKANLIFANGSTQAEFGTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 233

QY 76 IRFAGKEKTPPEE-----PKEEVTKANLIYADGKTQTAEFGKTFE 115  
Db 234 IFAGKEATNRNTDGSYDYLQINSRWGLTSAAEVTKANLIFANGSTQAEFGKTFE 293

QY 116 EATAEAYRYADALKKONGEYTVDVADKGYTLNLFKAGKEKTPPEPK 161  
Db 294 KATSEAYAYADTLKKDNGEYTVDVADKGYTLNLFKAGKESAWRHPQ 339

RESULT 12  
US-09-710-299-6  
Sequence 6, Application US/09710299  
Patent No. 6521741  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
Suess, Gabriele M.  
Tarlinton, David M.  
Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
PRODUCING SAME  
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/710,299  
FILING DATE: 09-No. 6521741-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-710-299-6

Query Match 27.7%; Score 618; DB 4; Length 342;  
Best Local Similarity 75.9%; Pred. No. 1.5e-41;  
Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;

QY 16 SEEEVTKANLIFANGSTQAEFGTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 75  
Db 174 SAEVTKANLIFANGSTQAEFGTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLN 233

QY 76 IRFAGKEKTPPEE-----PKEEVTKANLIYADGKTQTAEFGKTFE 115  
Db 234 IFAGKEATNRNTDGSYDYLQINSRWGLTSAAEVTKANLIFANGSTQAEFGKTFE 293

QY 116 EATAEAYRYADALKKONGEYTVDVADKGYTLNLFKAGKEKTPPEPK 161  
Db 294 KATSEAYAYADTLKKDNGEYTVDVADKGYTLNLFKAGKESAWRHPQ 339

RESULT 13  
US-08-378-761A-19  
Sequence 19, Application US/08378761A  
Patent No. 5635384  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
TITLE OF INVENTION: USING  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:



GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
APPLICANT: Suess, Gabriele M.  
APPLICANT: Tarlington, David M.  
APPLICANT: Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
TITLE OF INVENTION: PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,741B  
FILING DATE: 26-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-828-741B-13

Query Match 27.5%; Score 615; DB 3; Length 178;  
Best Local Similarity 76.2%; Pred. No. 1.le-41;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;  
QY 18 EEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKNNGEYTVDVADKGYTLNIK 77  
Db 12 EEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKNNGEYTVDVADKGYTLNIK 71  
QY 78 FAGKEKTPPE-----PKEVTTKANLIYADGKTQTAFFKGTPEEA 117  
Db 72 FAGKEATNRRTDGTGYILQINSRWGGLTSAEEVTTKANLIFANGSTQTAFFKGTPEKA 131  
QY 118 TAEAYRYADALKKNGEYTVDVADKGYTLNIKFPAGKEKTPPEPK 161  
Db 132 TSEAYAYADTLKKNNGEYTVDVADKGYTLNIKFPAGKESAWRHPQ 175

Search completed: September 3, 2003, 11:04:46  
Job time : 17.4438 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:03:32 ; Search time 45.2206 Seconds  
(without alignments)  
1317.758 Million cell updates/sec

Title: US-08-325-278B-3

Perfect score: 2235

Sequence: 1 AVENKEETPETPTDSEEV.....GVDGWTVDATKTFVTVM 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	100.0	434	8	US-08-325-278-3
2	1585	70.0	305	8	US-08-325-278-1
3	697.5	31.2	669	9	US-09-878-756-4
4	618	27.7	342	12	US-10-345-618-6
5	615	27.5	178	12	US-10-345-618-13
6	615	27.5	198	12	US-10-345-618-8
7	614	27.5	495	12	US-10-345-618-4
8	613	27.4	182	12	US-10-345-618-2
9	608	27.2	482	12	US-10-345-618-16
10	389	17.4	76	10	US-09-808-212A-4
11	371	16.6	72	10	US-09-808-212A-6
12	370	16.6	72	10	US-09-808-212A-10
13	369	16.5	72	10	US-09-808-212A-8
14	341	15.3	82	10	US-09-808-212A-2
15	336	15.0	166	12	US-10-338-411-21

Sequence 18, Appl  
Sequence 14, Appl  
Sequence 3, Appl  
Sequence 16, Appl  
Sequence 12, Appl  
Sequence 56, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 1, Appl  
Sequence 68, Appl  
Sequence 45, Appl  
Sequence 45, Appl  
Sequence 82, Appl  
Sequence 33, Appl  
Sequence 118, Appl  
Sequence 16, Appl  
Sequence 56, Appl  
Sequence 58, Appl  
Sequence 8, Appl  
Sequence 14, Appl  
Sequence 35, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 7, Appl  
Sequence 184, Appl  
Sequence 5, Appl  
Sequence 383, Appl  
Sequence 2, Appl  
Sequence 10, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-325-278-3  
; Sequence 3, Application US/08325278  
; Publication No. US20030027283A1  
; GENERAL INFORMATION:  
; APPLICANT: Bjvrck, Lars  
; APPLICANT: Sjbvring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278  
; FILING DATE: 26-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 450023.401  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown



```
Db 203 QAVOLEAPTVIDAPELTPALTYTKLVKVGNTF--SGETTK-----AIDTATAEKEFKQYA 256
Qy 126 DALKDNGETVDVADKGYTLNIKFA--GKEKTPPEPKBEVTIKANLIYADGKTQTAEF 182
Db 257 TANVVD--GWSYDDATKTFVTEKPAVIDALELTPALTYTKLVKGNF--SGETTKAI 313
Qy 183 KGTFEATAYAYADLLAKENGYTVVDVADKGYTLNIKFA---GKEKTPPEPKBEVTIK 239
Db 314 DAATAEKEFKQYATAN--NVDGWSYDYATKTFVTEKPAVIDAPELTPALTYTKLVK 370
Qy 240 ANLIYADGKTQTAEFKTFEATAYAYADLLAKENGYTDADLEDGYYINIRFAKKV 299
Db 371 GNTF--SGETTKAIDATAEKEFKQYATAN---NVDGWSYDDATKTFVTEKPA---V 422
Qy 300 DEKPE--EPMDTYKLIINGTKLGETTEAADAATAEKFVKQYANDGVGGEWYDDATK 357
Db 423 IDAPELTPALTYTKLVKGNFSGETTKAVDAETAEKAFAKQYATANNVDSYDDATK 482
Qy 358 TFWTEPEVIDASELTPAVTYTKLVINGTKLGETTKAVDAETAEKAFAKQYANDGV 417
Db 483 TFWTEKPAVIDAPELTPALTYTKLVINGTKLGETTKAVDVETAEKAFAKQYAN 542
Qy 418 GVWYDDATKTFVTTEM 434
Db 543 GVWYDDATKTFVTTEM 559

RESULT 4
US-10-345-618-6
; Sequence 6, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDS AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TLHL protein
; OTHER INFORMATION: sequence
US-10-345-618-6

Query Match 27.7%; Score 618; DB 12; Length 342;
Best Local Similarity 75.9%; Pred. No. 3.3e-38;
Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;

Qy 16 SEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVVDVADKGYTLN 75
Db 174 SAEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVVDVADKGYTLN 233
Qy 76 IKFAGKEKTPPE-----PKEEVTIKANLIYADGKTQTAEFKGTPE 115
Db 234 IKFAGKEATNRNTDGSYDGIQINSRWGLTSAEEVTIKANLIFANGSTQTAEFKGTPE 293
Qy 116 EATAEAYRYADALKDNGEYTVVDVADKGYTLNKFAGKEKTPPEPK 161
Db 294 KATSEAYAYADTLKKDNGEYTVVDVADKGYTLNKFAGKESAWRHPQ 339

RESULT 5
US-10-345-618-13
; Sequence 13, Application US/10345618
```

```
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDS AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LHL-omp
; OTHER INFORMATION: protein sequence
US-10-345-618-13

Query Match 27.5%; Score 615; DB 12; Length 178;
Best Local Similarity 76.2%; Pred. No. 2.3e-38;
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

Qy 18 EEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVVDVADKGYTLNIK 77
Db 12 EEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVVDVADKGYTLNIK 71
Qy 78 FAGKEKTPPE-----PKEEVTIKANLIYADGKTQTAEFKGTPEEA 117
Db 72 FAGKEATNRNTDGSYDGIQINSRWGLTSAEEVTIKANLIFANGSTQTAEFKGTPEKA 131
Qy 118 TAEAYRYADALKDNGEYTVVDVADKGYTLNKFAGKEKTPPEPK 161
Db 132 TSEAYAYADTLKKDNGEYTVVDVADKGYTLNKFAGKESAWRHPQ 175

RESULT 6
US-10-345-618-8
; Sequence 8, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDS AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LHL.seq
; OTHER INFORMATION: protein sequence
US-10-345-618-8

Query Match 27.5%; Score 615; DB 12; Length 198;
Best Local Similarity 76.2%; Pred. No. 2.6e-38;
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

Qy 18 EEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVVDVADKGYTLNIK 77
Db 32 EEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVVDVADKGYTLNIK 91
```

Qy	78	FAGKEKTPEE-----PKEVTKANLIYADOKTQTAEFKGTFEBA
Db	92	FAGKANNTDGTGYILOINRWGLTSAEVTIKANLIFANGSTQAEFKGTFEKA
Qy	118	TAEAYRYADALKKDNGEYTVDVADKGTYLNKFAGKEKTPPEPK
		:
Db	152	TSEAYAYADTLKKDNGEYTVDVADKGTYLNKFAGKESAWRHQ
		:

```

RESULT 7
US-10-345-618-4
  SEQUENCE 4, Application US/10345618
  Publication No. US20030148484A1
  GENERAL INFORMATION:
  APPLICANT: Koentgen, Frank
  APPLICANT: Suess, Gabriele M.
  APPLICANT: Tarlinton, David M.
  APPLICANT: Treutlein, Herbert R.
  TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
  FILE REFERENCE: 13474
  CURRENT APPLICATION NUMBER: US/10/345,618
  CURRENT FILING DATE: 2003-01-16
  PRIOR APPLICATION NUMBER: US/09/509,031
  PRIOR FILING DATE: 2000-06-09
  NUMBER OF SEQ ID NOS: 16
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 4
  LENGTH: 495
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence:CATAB-TEV
  OTHER INFORMATION: protein sequence
US-10-345-618-4

```

Query Match	27.5%	Score 614;	DB 12;	Length 495;
Best Local Similarity	78.6%;	Pred. No. 1.1e-37;		
Matches 125; Conservative 4;	Mismatches 10;	Indels 20;	Gaps 1;	
QY	16	SEEEVTKANLIFANGSTQTAEEKGTPEKATSEAYAYADTLKKONGEYTVDVADKGYTLN	75	
Db	174	SAEEVTKANLIFANGSTQTAEEKGTPEKATSEAYAYADTLKKONGEYTVDVADKGYTLN	233	
QY	76	IKFAGKEKTPPE-----PREEVTIKANLIYADGKTQTAEEKGTTFE	115	
Db	234	IKFAGKEATNRNTDGDYGLQLINRWGGUITSAAEVTIKANLIFANGSTQTAEEKGTTFE	293	
QY	116	EATAEAYRYADALKKONGEYTVDVADKGYTLNINIKFAGKE	154	
Db	294	KATSEAYAYADTLKKONGEYTVDVADKGYTLNINIKFAGKE	332	

```

RESULT 8
US-10-345-618-2
; Sequence 2, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 182
; TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Description of Artificial Sequence:LHL protein
;
; OTHER INFORMATION: sequence
US-10-345-618-2

```

	Query Match	27.4%	Score 613;	DB 12;	Length 182;
	Best Local Similarity	78.5%;	Pred. No. 3.3e-38;		
	Matches 124;	Conservative 5;	Mismatches 9;	Indels 20;	Gaps 1;
Qy	15 DSEEVYIKANLIFANGSTQTAEFGTFEKATSEAYADTLKKDNGEYTVDVADKGYTL	74	:		
Db	25 DNTEEVYIKANLIFANGSTQTAEFGTFEKATSEAYADTLKKDNGEYTVDVADKGYTL	84	:		
Qy	75 NIFAGKEKTPEEP-----KEEVTIKANLIYADKGTOAEFKGTF	114	:		
Db	85 NIFAGKEATNRWTDGGSDYGLIQLNSRWGLTLLKEVTIKANLIFANGSTQTAEFKGTF	144	:		
Qy	115 EETAEEAIRYADALKKDNGEYTVDVADKGYTLNLTKFAG	152	:		
Db	145 EKATSEAYADTLKKDNGEYTVDVADKGYTLNLTKFAG	182	:		

```

RESULT 9
US-10-345-618-16
; Sequence 16, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: comTlgL protein
; OTHER INFORMATION: sequence
US-10-345-618-16

```

Query Match	27.28;	Score 608;	DB 12;	Length 482;
Best Local Similarity	80.0%;	Pred. No. 2.9e-37;		
Matches 124;	Conservative 5;	Mismatches 10;	Indels 16;	Gaps 1;
Qy	16	SEEEVTKANLIIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVDVADKGTYTLN	75	
Dd	321	SAEEVTIKANLIIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYTVDVADKGTYTLN	380	
Qy	76	IKFAGKEKT-----PEEPKEEVTIKANLIIVADGKTTOTAEFFKGTFEATA	119	
	:			:
Dd	381	IKFAGKEASGGSGGGSGGGSGGGSGGAEEVTIKANLIIFANGSTQTAEFKGTFEKATS	440	
Qy	120	EAYRYADALKKDNGEYTVDVADKGYTTLNIKFAGKE	154	
Dd	441	EAYAYADTLKKDNGEYTVDVADKGYTTLNIKFAGKE	475	

RESULT 10  
US-09-808-212A-4  
; Sequence 4, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann

; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-4

Query Match 17.4%; Score 389; DB 10; Length 76;  
Best Local Similarity 100.0%; Pred. No. 4.1e-22;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KEETPTETDSEEVYTIKANLIFANGSTQTAEFKGTFAEATAYAYADTLKKNGEYT 64  
DB 1 KEETPTETDSEEVYTIKANLIFANGSTQTAEFKGTFAEATAYAYADTLKKNGEYT 60  
QY 65 VDVAADKGYTLNIRFAG 80  
DB 61 VDVAADKGYTLNIRFAG 76

RESULT 11  
US-09-808-212A-6  
; Sequence 6, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-6

Query Match 16.6%; Score 371; DB 10; Length 72;  
Best Local Similarity 100.0%; Pred. No. 8.2e-21;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 81 KEETPEEPKEEVYTIKANLIYADGKTQTAEFKGTFAEATAYADALKKDNGEYTVDA 140  
DB 1 KEETPEEPKEEVYTIKANLIYADGKTQTAEFKGTFAEATAYADALKKDNGEYTVDA 60  
QY 141 DKGYTLNIRFAG 152  
DB 61 DKGYTLNIRFAG 72

RESULT 12  
US-09-808-212A-10  
; Sequence 10, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-10

Query Match 16.6%; Score 370; DB 10; Length 72;  
Best Local Similarity 100.0%; Pred. No. 9.7e-21;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 225 KEETPEEPKEEVYTIKANLIYADGKTQTAEFKGTFAEATAYADLLAKENGYTVADLE 284  
DB 1 KEETPEEPKEEVYTIKANLIYADGKTQTAEFKGTFAEATAYADLLAKENGYTVADLE 60  
QY 285 DGGYTLNIRFAG 296  
DB 61 DGGYTLNIRFAG 72

RESULT 13  
US-09-808-212A-8  
; Sequence 8, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-8

Query Match 16.5%; Score 369; DB 10; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 153 KEETPEEPKEEVYTIKANLIYADGKTQTAEFKGTFAEATAYADLLAKENGYTVDA 212  
DB 1 KEETPEEPKEEVYTIKANLIYADGKTQTAEFKGTFAEATAYADLLAKENGYTVDA 60  
QY 213 DKGYTLNIRFAG 224  
DB 61 DKGYTLNIRFAG 72

RESULT 14  
US-09-808-212A-2  
; Sequence 2, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 82  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-2

Query Match 15.3%; Score 341; DB 10; Length 82;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:01:51 ; Search time 317.131 Seconds  
(without alignments)  
1191.101 Million cell updates/sec

Title: US-08-325-278B-3  
Perfect score: 2235  
Sequence: 1 AVENKEPPEPTDSEEV.....GVDGWVYDDATKFTVTM 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :			
Pending Patents_AA_Main:*			
1:	/cgn2_6/ptodata/1/paa/US06_COMB.pcp.*		
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pcp.*		
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pcp.*		
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
5:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
6:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
7:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
8:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
9:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
10:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
11:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
12:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
13:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
14:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*		
15:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
16:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
17:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
18:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
19:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
20:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
21:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
22:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
23:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
24:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
25:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*		
26:	/cgn2_6/ptodata/1/paa/US10_COMB.pcp.*		
27:	/cgn2_6/ptodata/1/paa/US10_COMB.pcp.*		
28:	/cgn2_6/ptodata/1/paa/US10_COMB.pcp.*		
29:	/cgn2_6/ptodata/1/paa/US10_COMB.pcp.*		
30:	/cgn2_6/ptodata/1/paa/US10_COMB.pcp.*		
31:	/cgn2_6/ptodata/1/paa/US60_COMB.pcp.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	100.0	434	7	US-08-325-278-3 Sequence 3, Appli
2	2235	100.0	434	7	US-08-325-278A-3 Sequence 3, Appli

3	2235	100.0	434	7	US-08-325-278B-3 Sequence 3, Appli
4	1587.5	71.0	719	22	US-09-791-537-10210 Sequence 10210, A
5	1587.5	71.0	719	22	US-09-791-537-96101 Sequence 96101, A
6	1565	70.0	305	7	US-08-325-278-1 Sequence 1, Appli
7	1565	70.0	305	7	US-08-325-278A-1 Sequence 1, Appli
8	1565	70.0	305	7	US-08-325-278B-1 Sequence 1, Appli
9	1372	69.4	467	25	US-09-980-469-12 Sequence 12, Appli
10	1372	69.4	467	25	US-09-980-469-12 Sequence 12, Appli
11	1263.5	56.5	992	22	US-09-791-537-88366 Sequence 88366, A
12	1263.5	56.5	1027	8	US-08-331-637-2 Sequence 2, Appli
13	1263.5	56.5	1027	8	US-08-446-137A-2 Sequence 2, Appli
14	1263.5	56.5	1027	15	US-09-187-295-2 Sequence 2, Appli
15	1216	54.4	291	8	US-08-446-137A-4 Sequence 4, Appli
16	708	31.7	593	22	US-09-791-537-86905 Sequence 86905, A
17	706.5	31.6	664	10	US-08-669-408-2 Sequence 2, Appli
18	706.5	31.6	664	10	US-08-669-408A-2 Sequence 2, Appli
19	697.5	31.2	669	23	US-09-878-756-4 Sequence 4, Appli
20	692.5	31.0	480	22	US-09-791-537-79119 Sequence 79119, A
21	683	30.6	448	22	US-09-791-537-86901 Sequence 86901, A
22	683	30.6	448	31	US-60-388-059-5 Sequence 5, Appli
23	670	30.0	185	17	US-09-315-208-2 Sequence 2, Appli
24	670	30.0	185	22	US-09-791-537-5341 Sequence 5341, Ap
25	634.5	28.4	228	23	US-09-889-182A-2 Sequence 2, Appli
26	627	28.1	250	23	US-09-889-182A-6 Sequence 6, Appli
27	625	28.0	208	5	US-08-110-653-11 Sequence 11, Appli
28	622.5	27.9	502	7	US-08-378-761-25 Sequence 25, Appli
29	622	27.8	493	7	US-08-378-761-23 Sequence 23, Appli
30	618	27.7	342	19	US-09-509-031-6 Sequence 6, Appli
31	618	27.7	342	23	US-09-820-048A-6 Sequence 6, Appli
32	618	27.7	489	7	US-08-378-761-19 Sequence 19, Appli
33	615	27.5	178	19	US-09-509-031-13 Sequence 13, Appli
34	615	27.5	178	23	US-09-820-048A-13 Sequence 13, Appli
35	615	27.5	198	23	US-09-509-031-8 Sequence 8, Appli
36	615	27.5	198	23	US-09-820-048A-8 Sequence 8, Appli
37	614	27.5	495	19	US-09-509-031-4 Sequence 4, Appli
38	614	27.5	495	23	US-09-820-048A-4 Sequence 4, Appli
39	613	27.4	182	19	US-09-509-031-2 Sequence 2, Appli
40	613	27.4	182	23	US-09-820-048A-2 Sequence 2, Appli
41	608	27.2	482	19	US-09-509-031-16 Sequence 16, Appli
42	481	21.5	429	22	US-09-791-537-146174 Sequence 146174, A
43	404	18.1	413	10	US-08-669-408-10 Sequence 10, Appli
44	404	18.1	413	10	US-08-669-408A-10 Sequence 10, Appli
45	404	18.1	413	22	US-09-791-537-131407 Sequence 131407, A

ALIGNMENTS

RESULT 1  
US-08-325-278-3  
; Sequence 3, Application US/08325278  
; GENERAL INFORMATION:  
; APPLICANT: Bjvrck, Lars  
; APPLICANT: Sjvbring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278  
; FILING DATE: 26-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055
; US-08-325-278-3

Query Match 100.0%; Score 2235; DB 7; Length 434;
Best Local Similarity 100.0%; Pred. No. 7.5e-162;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEEVITKANLIIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
DB 1 AVENKEETPETDSEEEVITKANLIIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
QY 61 GEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTAEFKGTPEEATAE 120
DB 61 GEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTAEFKGTPEEATAE 120
QY 121 AYRYADALKKNGEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTA 180
DB 121 AYRYADALKKNGEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTA 180
QY 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKA 240
DB 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKA 240
QY 241 NLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGYTVADLEGGYTTINIRFAGKVD 300
DB 241 NLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGYTVADLEGGYTTINIRFAGKVD 300
QY 301 EKPEPMDTYKLIILNGKTLKGETTTAEVDAATAEKFVKQYANDNGVDGWTYDDATKTFT 360
DB 301 EKPEPMDTYKLIILNGKTLKGETTTAEVDAATAEKFVKQYANDNGVDGWTYDDATKTFT 360
QY 361 VTEKPEVIDASELTPAVTTYKLIVNGKTLKGETTTKAVDAETAFAKAFQYANDNGVDGVW 420
DB 361 VTEKPEVIDASELTPAVTTYKLIVNGKTLKGETTTKAVDAETAFAKAFQYANDNGVDGVW 420
QY 421 TYDDATKTFTVTTEM 434
DB 421 TYDDATKTFTVTTEM 434

RESULT 2
US-08-325-278A-3
; Sequence 3, Application US/08325278A
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICANT INFORMATION:
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
; US-08-325-278A-3

Query Match 100.0%; Score 2235; DB 7; Length 434;
Best Local Similarity 100.0%; Pred. No. 7.5e-162;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEEVITKANLIIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
DB 1 AVENKEETPETDSEEEVITKANLIIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
QY 61 GEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTAEFKGTPEEATAE 120
DB 61 GEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTAEFKGTPEEATAE 120
QY 121 AYRYADALKKNGEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTA 180
DB 121 AYRYADALKKNGEYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKANLIYADGKTQTA 180
QY 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKA 240
DB 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNKFAGKEKTPPEEKKEEVTIKA 240
QY 241 NLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGYTVADLEGGYTTINIRFAGKVD 300
DB 241 NLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGYTVADLEGGYTTINIRFAGKVD 300
QY 301 EKPEPMDTYKLIILNGKTLKGETTTAEVDAATAEKFVKQYANDNGVDGWTYDDATKTFT 360
DB 301 EKPEPMDTYKLIILNGKTLKGETTTAEVDAATAEKFVKQYANDNGVDGWTYDDATKTFT 360
QY 361 VTEKPEVIDASELTPAVTTYKLIVNGKTLKGETTTKAVDAETAFAKAFQYANDNGVDGVW 420
DB 361 VTEKPEVIDASELTPAVTTYKLIVNGKTLKGETTTKAVDAETAFAKAFQYANDNGVDGVW 420
QY 421 TYDDATKTFTVTTEM 434
DB 421 TYDDATKTFTVTTEM 434

RESULT 3
US-08-325-278B-3
; Sequence 3, Application US/08325278B
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group

```





US-09-791-537-96101

```
Query Match          71.0%; Score 1587.5; DB 22; Length 719;
Best Local Similarity 76.3%; Pred. No. 4.6e-112;
Matches 334; Conservative 13; Mismatches 56; Indels 35; Gaps 6;

QY 2 VENKEETPETDSEEEVTKANLIFANGSTQTAEFKGTFFKATSEAYAYADTLKKDNG 61
   |||||
Db 95 VENKEETPETDSEEEVTKANLIFANGSTQTAEFKGTFFKATSEAYAYADTLKKDNG 154

QY 62 EYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFFKATAEA 121
   |||||
Db 155 EYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFFKATAEA 214

QY 122 YRYADALKKDNGEYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 181
   |||||
Db 215 YRYADALKKDNGEYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 274

QY 182 FKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKAN 241
   |||||
Db 275 FKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKAN 334

QY 242 LIYADGKTQTAEFKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEKTP 301
   |||||
Db 335 LIYADGKTQTAEFKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEKTP 394

QY 302 KPEPMD-TYK--LILNGKTLKGETTEAATAEKVFKQYANDNGVD-GEWYDDATK 357
   |||||
Db 395 KPEKQVTKENIYEPEDGVQTATPKGTFAEATAEAY--RYADLLSKHEKGYTADLDG 452

QY 358 TETVT-----EKPEVIDASELTPATVTVKLVINGKTLKGETTKAVDAETA 404
   |||||
Db 453 GYTNIRFAGKEPETPEKPEVQD-----GYASYEAEAAAKEALKND 496

QY 405 KAFKQYANDNGVDGVWY 422
   |||||
Db 497 DVNKSYYTIRQAGRYYY 514
```

RESULT 6

```
US-08-325-278-1
; Sequence 1, Application US/08325278
; GENERAL INFORMATION:
; APPLICANT: Bjvrck, Lars
; APPLICANT: Sjobring, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMASTERS, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
```

RESULT 7

```
US-08-325-278A-1
; Sequence 1, Application US/08325278A
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; APPLICANT: Sjobring, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278A
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
```

```
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; US-08-325-278-1

Query Match          70.0%; Score 1565; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.7e-111;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEEVTKANLIFANGSTQTAEFKGTFFKATSEAYAYADTLKKDN 60
   |||||
Db 1 AVENKEETPETDSEEEVTKANLIFANGSTQTAEFKGTFFKATSEAYAYADTLKKDN 60

QY 61 GEYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFFKATAE 120
   |||||
Db 61 GEYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFFKATAE 120

QY 121 AYRYADALKKDNGEYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180
   |||||
Db 121 AYRYADALKKDNGEYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180

QY 181 EFKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKA 240
   |||||
Db 181 EFKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEKTPPEPKKEVTIKA 240

QY 241 NLIYADGKTQTAEFKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEK 300
   |||||
Db 241 NLIYADGKTQTAEFKGTFFKATAEAAYRYADLLAKENGYTVADVADGYTLNIFKAGKEK 300

QY 301 EKPEE 305
   |||||
Db 301 EKPEE 305
```

```
;
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278A-1

Query Match      70.0%; Score 1565; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.7e-111;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
   |||||
Db 1 AVENKEETPETPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60

QY 61 GEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTAEFGTFEEATAE 120
   |||||
Db 61 GEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTAEFGTFEEATAE 120

QY 121 AYRYADALKKDNGEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTA 180
   |||||
Db 121 AYRYADALKKDNGEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTA 180

QY 181 EFKGTPEEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKA 240
   |||||
Db 181 EFKGTPEEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKA 240

QY 241 NLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEK 300
   |||||
Db 241 NLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEK 300

QY 301 EKPEE 305
   |||||
Db 301 EKPEE 305

RESULT 8
US-08-325-278B-1
; Sequence 1, Application US/08325278B
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278B
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
```

```
;
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278B-1

Query Match      70.0%; Score 1565; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.7e-111;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
   |||||
Db 1 AVENKEETPETPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60

QY 61 GEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTAEFGTFEEATAE 120
   |||||
Db 61 GEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTAEFGTFEEATAE 120

QY 121 AYRYADALKKDNGEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTA 180
   |||||
Db 121 AYRYADALKKDNGEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTA 180

QY 181 EFKGTPEEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKA 240
   |||||
Db 181 EFKGTPEEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKA 240

QY 241 NLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEK 300
   |||||
Db 241 NLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKGYTVADVADKGYTLNKFAGKEK 300

QY 301 EKPEE 305
   |||||
Db 301 EKPEE 305

RESULT 9
US-09-980-469-12
; Sequence 12, Application US/09980469
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND R
; FILE OF INVENTION: PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTUR
; FILE REFERENCE: 01/22924
; CURRENT APPLICATION NUMBER: US/09/980,469
; CURRENT FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: pUC19-cell-protL-cekNG-HDEL fusion encoded product
US-09-980-469-12

Query Match      69.4%; Score 1550; DB 25; Length 467;
Best Local Similarity 79.2%; Pred. No. 1.8e-109;
Matches 328; Conservative 3; Mismatches 43; Indels 40; Gaps 7;

QY 1 AVENKEETPETPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 60
   |||||
Db 40 AVENKEETPETPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDN 99

QY 61 GEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTAEFGTFEEATAE 120
   |||||
Db 100 GEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTAEFGTFEEATAE 159

QY 121 AYRYADALKKDNGEYTVADVADKGYTLNKFAGKEKTPPEPKKEEVIKANLIYADGKTQTA 180
   |||||
```

```
Db 160 AYRYADALKKNGEYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIYADGKTQTA 219
QY 181 EFKGTFEEATAEAYRYADLL-ARENGKYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIK 239
Db 220 EFKGTFEEATAEAYRYADLLAAKENGKYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIK 279
QY 240 ANLIYADGKTQTAEFKGTAEATAEAYRYADLLAKENGKYTDADLEGGYTTINIRFAGKV 299
Db 280 ANLIYADGKTQTAEFKGTAEATAEAYRYADLLAKENGKYTDADLEGGYTTINIRFAGKV 339
QY 300 DEKPEPMDYKILNKGKTLKGETTTEAYDAA-----TAEKVFQYANDN 344
Db 340 DEKPEGIPT-----PTPSASGPGCCQVLWGWNQWNTGFTAQVTVKN-TGSA 386
QY 345 GVDGEWYDDAKYTFVTTEKPEVIDASELT-----PAVTYKLVINGKTLKGET 393
Db 387 PVDG-WTL-----TFSFPGQQVQTQAWSSTVTSQSAVTVRNAPWNGNPAGGT 434

RESULT 10
US-09-889-182A-4
; Sequence 4, Application US/09889182A
; GENERAL INFORMATION:
; APPLICANT: Breitling, Frank
; TITLE OF INVENTION: SELECTION OF MONOCLONAL ANTIBODIES
; FILE REFERENCE: 4121-126
; CURRENT APPLICATION NUMBER: US/09/889,182A
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: pct/de00/00079
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-889-182A-4

Query Match 61.4%; Score 1372; DB 23; Length 367;
Best Local Similarity 82.4%; Pred. No. 4.9e-96;
Matches 271; Conservative 19; Mismatches 29; Indels 10; Gaps 3;

QY 5 KEETPTPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNGEYT 64
Db 25 KEKTPPEP---KEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADALKKNGEYT 80
QY 65 VDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRY 124
Db 81 VDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRY 140
QY 125 ADALKKNGEYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIYADGKTQTAEFK 184
Db 141 ADALKKNGEYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIYADGKTQTAEFK 200
QY 185 TFEATAEAYRYADLLAKENGKYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIY 244
Db 201 TFEATAEAYRYADALKKNGEYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIY 260
QY 245 ADGKTQTAEFKGTAEATAEAYRYADLLAKENGKYTDADLEGGYTTINIRFAGKVDKPE 304
Db 261 ADGKTQTAEFKGTAEATAEAYRYADALKKNGEYTVDVADKGYTLNIRKFAKKEKTPPE 320
QY 305 EPMDYKILNKGKTLKGETTTEAVDAATA 333
Db 321 SEED-----LNG-AVDGQNDTSTSPSA 343

RESULT 11
US-09-791-537-88366
; Sequence 88366, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88366
; LENGTH: 992
; TYPE: PRT
; ORGANISM: Peptostreptococcus magnus
US-09-791-537-88366

Query Match 56.5%; Score 1263.5; DB 22; Length 992;
Best Local Similarity 58.8%; Pred. No. 4.2e-87;
Matches 281; Conservative 36; Mismatches 90; Indels 71; Gaps 12;

QY 2 VENKEETPTPTDSEEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKKDNG 61
Db 241 MERKLESEKTEPE--PEEEVTIKANLIFADGSTQNAEFKGTFAKAVSDAYADALKKNG 298
QY 62 EYTVDVADKGYTLNIRKFAKKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTPEEATAEA 121
Db 299 EYTVDVADKGLTLNIRKFAKKEKTPPEPKKEVTIKVNLIFADGKTQTAEFKGTPEEATAKA 358
QY 122 YRYADALKKNGEYTVDVADKGYTLNIRKFAKKE--KTPPEPKKEVTIKANLIYADGKTOT 179
Db 359 YAYADLLAKENGKYTDADLEGGYTTINIRKFAKKEKTPPEPKKEVTIKVNLIFADGKIQT 418
QY 180 AEFKGTPEEATAEAYRYADLLAKENGKYTVDVADKGYTLNIRKFAKKE--KTPPEPKKEVT 237
Db 419 AEFKGTPEEATAKAYAYANLLAKENGKYTDADLEGGYTTINIRKFAKKEKTPPEPKKEVT 478
QY 238 IKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDADLEGGYTTINIRFAGK 297
Db 479 IKVNLIFADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKYTDADLEGGYTTINIRFAGK 538
QY 298 K-----VDE-----KPEEPM-----DTY-KLILNGKTLKG----- 321
Db 539 EQGEPNGPITIDELWLLKNAKEAEIKELKEAGITSDFSLINKAKTVEGVEALKNELKA 598
QY 322 ----ET-----TTEAVDAATAEAKVFQYAND-----NGVDGEWYDDATFTFTVTEK 364
Db 599 HAGEETPELKDGATYEEAEAAKALKNDVDVNNAYEIVQGADGRYYY--VLKIEVADEE 656
QY 365 PEVIDASELTPAVTVTKLVINGKTLKGETTTRKAVDAETAETAKPKQYANDNGVDGVWY 422
Db 657 EPGEDTPEVOEGYATYE-----EAEAAAKEALKEDKVNNAVYEVQVQADGRYYY 704

RESULT 12
US-08-331-637-2
; Sequence 2, Application US/08331637
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEIN L AND PROCESS FOR ITS PREPARATION BY
; TITLE OF INVENTION: RECOMBINANT DNA TECHNOLOGY
; NUMBER OF SEQUENCES: 23
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

Query Match 56.5%; Score 1263.5; DB 8; Length 1027;

RESULT 14  
US-09-187

	Query Match	56.3%	Score 1263.5;	DB 15;	Length 1027;
	Best Local Similarity	58.8%;	Pred. No. 4.4e-87;		
	Matches 281;	Conservative 36;	Mismatches 90;	Indels 71;	Gaps 12;
QY	2	VENKEETPEPTDSREEVITKANLIFANGSTOTAEFGTFEKATSEAYAYADTLKKDNG	61		
	:	:	:	:	:
Db	276	WERKLESEKTEPE--PEEEVITKANLIFADGSTQNAEEFGTFAKAVSAYAYADLKKDNG	333		
QY	62	EYTVDVADKGYTLNIRFAGKEKTPPEEPKEEVTIKANLIYADGKTQTQAEFGTFFEEATAEA	121		
	:	:	:	:	:



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:02:46 ; Search time 17.6184 Seconds  
(without alignments)  
773.734 Million cell updates/sec

Title: US-08-325-278B-3  
Perfect score: 2235  
Sequence: 1 AVENKEETPETDSEEV.....GVGWTYDDATKFTVTTEM 434

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 148256 seqs, 31410045 residues

Total number of hits satisfying chosen parameters: 148256

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	683	30.6	448	6	US-10-460-524-5
2	287	12.8	55	1	PCT-US03-21559-3
3	287	12.8	55	1	PCT-US03-21559-5
4	286	12.8	58	6	US-10-458-297-4
5	285	12.8	60	6	US-10-458-297-2
6	284	12.7	55	1	PCT-US03-21559-4
7	281.5	12.6	58	6	US-10-458-297-1
8	170.5	7.6	1849	6	US-10-637-544-2
9	160	7.2	2705	1	PCT-US03-21510-97
10	160	7.2	2705	6	US-10-408-765A-2039
11	160	7.2	2705	6	US-10-089-320B-44
12	157.5	7.0	1166	5	US-09-200-650E-7
13	156	7.0	916	6	US-10-408-765A-1222
14	156	7.0	916	7	US-60-490-890-475
15	155.5	7.0	770	5	US-09-897-516A-4453
16	153.5	6.9	701	6	US-10-333-120A-7
17	148.5	6.6	2647	1	PCT-US03-21510-95
18	147	6.6	521	6	US-10-430-752A-5
19	145.5	6.5	2315	6	US-10-603-114-5434
20	144	6.4	728	6	US-10-467-534-81
21	140	6.3	693	6	US-10-333-120A-10
22	140	6.3	901	6	PCT-US03-21510-96
23	136.5	6.1	2602	1	US-10-603-113-14747
24	136	6.1	630	6	US-10-603-113-20275
25	134.5	6.0	1262	6	US-10-617-320-5067
26	133	6.0	458	6	US-10-467-421-37

27	133	6.0	1742	6	US-10-615-383-4	Sequence 4, Appl
28	132.5	5.9	2060	6	US-10-381-596A-2	Sequence 2, Appl
29	132	5.9	487	5	US-09-897-516A-5285	Sequence 5285, Ap
30	131	5.9	6669	1	PCT-US03-21379-80	Sequence 80, Appl
31	131	5.9	930	5	US-09-200-650E-3	Sequence 3, Appl
32	130	5.8	1277	6	US-10-326-956-665	Sequence 665, App
33	130	5.8	1965	6	US-10-640-833-3829	Sequence 3829, Ap
34	130	5.8	1972	6	US-10-617-320-3251	Sequence 3251, Ap
35	129.5	5.8	1204	6	US-10-640-833-4083	Sequence 4083, Ap
36	128	5.7	718	6	US-10-603-108-2753	Sequence 2753, Ap
37	126	5.6	542	5	US-09-820-843B-4	Sequence 4, Appl
38	125.5	5.6	1315	5	US-09-200-650E-5	Sequence 5, Appl
39	124	5.5	1899	6	US-10-631-467-919	Sequence 919, App
40	124	5.5	3063	6	US-10-631-467-918	Sequence 918, App
41	123.5	5.5	1359	6	US-10-326-956-1053	Sequence 1053, Ap
42	121.5	5.4	2142	6	US-10-603-108-3459	Sequence 3459, Ap
43	121	5.4	1531	1	PCT-US02-37235-44	Sequence 44, Appl
44	120.5	5.4	818	5	US-09-897-516A-5123	Sequence 5123, Ap
45	120.5	5.4	1939	6	US-10-408-765A-2188	Sequence 2188, Ap

ALIGNMENTS

RESULT 1  
US-10-460-524-5  
; Sequence 5, Application US/10460524  
; GENERAL INFORMATION:  
; APPLICANT: Hernan, Ronald A  
; APPLICANT: Mehig, Richard J  
; APPLICANT: Brookie, Ian  
; APPLICANT: Jenkins, Elizabeth  
; TITLE OF INVENTION: Affinity Peptides and Method for Purification of Recombinant P  
; FILE REFERENCE: SGM 7047.1  
; CURRENT APPLICATION NUMBER: US/10/460,524  
; CURRENT FILING DATE: 2003-06-12  
; PRIOR APPLICATION NUMBER: US 60/388,059  
; PRIOR FILING DATE: 2002-06-12  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Streptococcus  
US-10-460-524-5

Query Match	30.6%;	Score 683;	DB 6;	Length 448;
Best Local Similarity	53.2%;	Pred. No. 4e-38;		
Matches 176;	Conservative 15;	Mismatches 92;	Indels 48;	Gaps 9;
QY	130	KDNGEYTVDAKGYTLNFKAGKEKTPPEPKEEVIKANLIYAD-GKTQAEFKGTTEE	188	
DB	45	RNGELT-----NLLGNSETTLALRNEESATADLTAAAVADTVAAAAENAGA	92	
QY	189	ATAEAYRYADLLA-----KENGKYTVDVADKGYTLNFKAGKEKTPPEPKEEVIKAN	241	
DB	93	AAWEAAAAADALAKAKADALKEFNKYGSVDYKLNINNAKTVEGIDKLAQVVSAAKAR	152	
QY	242	LIYA-DG-----KTQAEFKGTFAEATAEAYRYAD---LLAKENGKYTADLEDGGYTTINI	292	
DB	153	ISEATDGLSDFLKSQTP-----AEDTVKSIELAEAKVLANRELDKYGV----SDYHKNL	202	
QY	293	RFAGKKVDKPE-----BPMDFYKLLILNGKTLKGETTTTAEVDAATAEKFVKQYAND	343	
DB	203	INNAKTVEGVKELIDELIALPKTDYKLLILNGKTLKGETTTEAEVDAATAEKFVKQYAND	262	
QY	344	NGVDGEWYDDATKFTTVEKPEVIDASELTFAVTTYKLVINGKTLKGETTTTKAVDAETA	403	
DB	263	NGVDGEWYDDATKFTTVEKPEVIDASELTFAVTTYKLVINGKTLKGETTTTKAVDAETA	322	
QY	404	EKAFKQYANDNGVDGWTYDDATKFTVTTEM	434	
DB	323	EKAFKQYANDNGVDGWTYDDATKFTVTTEM	353	

```
RESULT 2
PCT-US03-21559-3
; Sequence 3, Application PC/TUS0321559
; GENERAL INFORMATION:
; APPLICANT: President Biosystems
; TITLE OF INVENTION: PROTEIN CHIPS
; FILE REFERENCE: 13744-002WO1
; CURRENT APPLICATION NUMBER: PCT/US03/21559
; CURRENT FILING DATE: 2003-07-10
; PRIOR APPLICATION NUMBER: US 10/193,377
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21559-3

Query Match      12.8%; Score 287; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 4.3e-13;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 TYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 363
DB 1 TYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 55

RESULT 3
PCT-US03-21559-5
; Sequence 5, Application PC/TUS0321559
; GENERAL INFORMATION:
; APPLICANT: President Biosystems
; TITLE OF INVENTION: PROTEIN CHIPS
; FILE REFERENCE: 13744-002WO1
; CURRENT APPLICATION NUMBER: PCT/US03/21559
; CURRENT FILING DATE: 2003-07-10
; PRIOR APPLICATION NUMBER: US 10/193,377
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21559-5

Query Match      12.8%; Score 287; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 4.3e-13;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 TYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 433
DB 1 TYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 55

RESULT 4
US-10-458-297-4
; Sequence 4, Application US/10458297
; GENERAL INFORMATION:
; APPLICANT: Eiichi Ogino
; APPLICANT: Michio Nomura
; APPLICANT: Takashi Asahi
; APPLICANT: Shuichi Kaneko
; APPLICANT: Akito Sakai
; TITLE OF INVENTION: ADSORBENT FOR ELIMINATING HEPATITIS C VIRUS, ADSORBENT, AND ADSOR
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/458,297
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: US/09/380,644C
```

```
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: JP 09/71483
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: word
; SEQ ID NO 4
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: lgg - binding peptide
US-10-458-297-4
```

```
Query Match      12.8%; Score 286; DB 6; Length 58;
Best Local Similarity 96.5%; Pred. No. 5.4e-13;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 377 VTTYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 433
DB 1 MTTYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 57
```

```
RESULT 5
US-10-458-297-2
; Sequence 2, Application US/10458297
; GENERAL INFORMATION:
; APPLICANT: Eiichi Ogino
; APPLICANT: Michio Nomura
; APPLICANT: Takashi Asahi
; APPLICANT: Shuichi Kaneko
; APPLICANT: Akito Sakai
; TITLE OF INVENTION: ADSORBENT FOR ELIMINATING HEPATITIS C VIRUS, ADSORBENT, AND ADS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/458,297
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: US/09/380,644C
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: JP 09/71483
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: word
; SEQ ID NO 2
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: lgg - binding peptide
US-10-458-297-2
```

```
Query Match      12.8%; Score 285; DB 6; Length 60;
Best Local Similarity 98.2%; Pred. No. 6.5e-13;
Matches 55; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 378 TTYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 433
DB 5 TTYKLINGKTLKGETTTKAVDAETAEKAFKQYANDNGVDGWTYDDATKTFVTVE 60
```

```
RESULT 6
PCT-US03-21559-4
; Sequence 4, Application PC/TUS0321559
; GENERAL INFORMATION:
; APPLICANT: President Biosystems
; TITLE OF INVENTION: PROTEIN CHIPS
; FILE REFERENCE: 13744-002WO1
; CURRENT APPLICATION NUMBER: PCT/US03/21559
; CURRENT FILING DATE: 2003-07-10
; PRIOR APPLICATION NUMBER: US 10/193,377
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
```



QY 73 TLNKFAGKEKTPEEPKKEVTIKANLIYADGKTQTAEFKGTFEEATAEA--YRYADALKK 130

QY 169 NLIYADGKTQTAEFKGTTEE-----ATAEAYRYADLLAKENGKTYVDVADKG 215  
DB 969 NTRGAGGQGO-LDVRMTSPSRPPIPKLEPGGAEAAQAVRY---MPPEGPYKVDITVDG 1024  
QY 216 YTL-NIKFAGKEKTEPEEKVEVT-----IKANLI-----YADGKTQTAEFKGTFAEATAE 264  
DB 1025 HPVPGSPFAVEGLPPDPSPKVCAYGPKLGGLVGTTPAPFSDTKAGTGGGLGTLVEGPCE 1084  
QY 265 AYRYADLLAKENGKTYADL-----EDGGYTINIRFAGKKVDEKP-----EEDMDTYKLIL 314  
DB 1085 ----AKIECQDNGDSCAVSLPTPEGTYTINILFAEAHIPGSPFKATIRPVDFDPSKVRA 1140  
QY 315 NGKTLK-----GETTTEAVDAATAEK-----VFKQYANDNGVDGEW--TYDDA- 355  
DB 1141 SGPGLERKVGAEATFTVDCSEAGEAELTIELSDAGVKAELVIHNNADGYHITYSPAF 1200  
QY 356 TKTFVTVEK-----PEVIDASELTPAVTYYKLVINGK-----TLKGETTTKAVDAET- 402  
DB 1201 PGTYYTITIKYGGHPVPKFPTRVHVQPAVDTSQVKVSGPVEPHGVLRVTTFTVTDARSL 1260  
QY 403 -----AEKAFQYANDNGVDGVM 420  
DB 1261 TATGNNHVTARVLNPSGAKTDITYVDNG-DGTY 1292

## RESULT 10

US-10-408-765A-2039  
; Sequence 2039, Application US/10408765A  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2039  
; LENGTH: 2705  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-2039

Query Match 7.2%; Score 160; DB 6; Length 2705;  
Best Local Similarity 22.2%; Pred. No. 0.0095;  
Matches 114; Conservative 60; Mismatches 187; Indels 152; Gaps 25;  
QY 47 SEAYAYADTLKKDNGEYTVDVADKG---YTLNKFAGKE---KTPEEPKEEYTIKANLIYA 101  
DB 793 READIDFDLIKNDNTFTVKYTPPGAGRYTIMLVFANQEIAPSPFIKVDPSHDSKVA 852  
QY 102 DGK-----TQTA-----EFKGTTEEAATAEAYRYADALKDNGEY 136  
DB 853 EGPGLNRTGVEYKPTHTFTVLTKGAGKAKLDVQFAGT---AKGEVVRDFEIIDNHDSYT 909  
QY 137 VD---VADKGYTLNKFAGKEKTPPEP-----KEEVTIKA 168  
DB 910 VKYTAVOQGNMAVTYTG-DVPKSPFVNVNAPPDLDSKIKVQGLNSKVAVGQQAQFSV 968  
QY 169 NLIYADGKTQTAEFKGTTEE-----ATAEAYRYADLLAKENGKTYVDVADKG 215  
DB 969 NTRGAGGQGO-LDVRMTSPSRPPIPKLEPGGAEAAQAVRY---MPPEGPYKVDITVDG 1024  
QY 216 YTL-NIKFAGKEKTEPEEKVEVT-----IKANLI-----YADGKTQTAEFKGTFAEATAE 264  
DB 1025 HPVPGSPFAVEGLPPDPSPKVCAYGPKLGGLVGTTPAPFSDTKAGTGGGLGTLVEGPCE 1084  
QY 137 VD---VADKGYTLNKFAGKEKTPPEP-----KEEVTIKA 168  
DB 910 VKYTAVOQGNMAVTYTG-DVPKSPFVNVNAPPDLDSKIKVQGLNSKVAVGQQAQFSV 968  
QY 169 NLIYADGKTQTAEFKGTTEE-----ATAEAYRYADLLAKENGKTYVDVADKG 215  
DB 969 NTRGAGGQGO-LDVRMTSPSRPPIPKLEPGGAEAAQAVRY---MPPEGPYKVDITVDG 1024  
QY 216 YTL-NIKFAGKEKTEPEEKVEVT-----IKANLI-----YADGKTQTAEFKGTFAEATAE 264  
DB 1025 HPVPGSPFAVEGLPPDPSPKVCAYGPKLGGLVGTTPAPFSDTKAGTGGGLGTLVEGPCE 1084

QY 265 AYRYADLLAKENGKTYADL-----EDGGYTINIRFAGKKVDEKP-----EEDMDTYKLIL 314  
DB 1085 ----AKIECQDNGDSCAVSLPTPEGTYTINILFAEAHIPGSPFKATIRPVDFDPSKVRA 1140  
QY 315 NGKTLK-----GETTTEAVDAATAEK-----VFKQYANDNGVDGEW--TYDDA- 355  
DB 1141 SGPGLERKVGAEATFTVDCSEAGEAELTIELSDAGVKAELVIHNNADGYHITYSPAF 1200  
QY 356 TKTFVTVEK-----PEVIDASELTPAVTYYKLVINGK-----TLKGETTTKAVDAET- 402  
DB 1201 PGTYYTITIKYGGHPVPKFPTRVHVQPAVDTSQVKVSGPVEPHGVLRVTTFTVTDARSL 1260  
QY 403 -----AEKAFQYANDNGVDGVM 420  
DB 1261 TATGNNHVTARVLNPSGAKTDITYVDNG-DGTY 1292

## RESULT 11

US-10-089-320B-44  
; Sequence 44, Application US/10089320B  
; GENERAL INFORMATION:  
; APPLICANT: KYOWA HAKKO KOGYO CO., LTD.  
; TITLE OF INVENTION: Shear Stress-Responsive Genes  
; FILE REFERENCE: 1241.22  
; CURRENT APPLICATION NUMBER: US/10/089,320B  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: JP 1999-280976  
; PRIOR FILING DATE: 1999-10-01  
; NUMBER OF SEQ ID NOS: 181  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 44  
; LENGTH: 2705  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-089-320B-44

Query Match 7.2%; Score 160; DB 6; Length 2705;  
Best Local Similarity 22.2%; Pred. No. 0.0095;  
Matches 114; Conservative 60; Mismatches 187; Indels 152; Gaps 25;  
QY 47 SEAYAYADTLKKDNGEYTVDVADKG---YTLNKFAGKE---KTPEEPKEEYTIKANLIYA 101  
DB 793 READIDFDLIKNDNTFTVKYTPPGAGRYTIMLVFANQEIAPSPFIKVDPSHDSKVA 852  
QY 102 DGK-----TQTA-----EFKGTTEEAATAEAYRYADALKDNGEY 136  
DB 853 EGPGLNRTGVEYKPTHTFTVLTKGAGKAKLDVQFAGT---AKGEVVRDFEIIDNHDSYT 909  
QY 137 VD---VADKGYTLNKFAGKEKTPPEP-----KEEVTIKA 168  
DB 910 VKYTAVOQGNMAVTYTG-DVPKSPFVNVNAPPDLDSKIKVQGLNSKVAVGQQAQFSV 968  
QY 169 NLIYADGKTQTAEFKGTTEE-----ATAEAYRYADLLAKENGKTYVDVADKG 215  
DB 969 NTRGAGGQGO-LDVRMTSPSRPPIPKLEPGGAEAAQAVRY---MPPEGPYKVDITVDG 1024  
QY 216 YTL-NIKFAGKEKTEPEEKVEVT-----IKANLI-----YADGKTQTAEFKGTFAEATAE 264  
DB 1025 HPVPGSPFAVEGLPPDPSPKVCAYGPKLGGLVGTTPAPFSDTKAGTGGGLGTLVEGPCE 1084  
QY 265 AYRYADLLAKENGKTYADL-----EDGGYTINIRFAGKKVDEKP-----EEDMDTYKLIL 314  
DB 1085 ----AKIECQDNGDSCAVSLPTPEGTYTINILFAEAHIPGSPFKATIRPVDFDPSKVRA 1140  
QY 315 NGKTLK-----GETTTEAVDAATAEK-----VFKQYANDNGVDGEW--TYDDA- 355  
DB 1141 SGPGLERKVGAEATFTVDCSEAGEAELTIELSDAGVKAELVIHNNADGYHITYSPAF 1200  
QY 356 TKTFVTVEK-----PEVIDASELTPAVTYYKLVINGK-----TLKGETTTKAVDAET- 402  
DB 1201 PGTYYTITIKYGGHPVPKFPTRVHVQPAVDTSQVKVSGPVEPHGVLRVTTFTVTDARSL 1260  
QY 403 -----AEKAFQYANDNGVDGVM 420

Db 1261 TATGGNHVTVARVLPNSGAKTDTYVTDNG-DGT 1292

## RESULT 12

US-09-200-650E-7

; Sequence 7, Application US/09200650E

; GENERAL INFORMATION:

; APPLICANT: Patti, Joseph M.

; APPLICANT: Foster, Timothy J.

; APPLICANT: Hook, Magnus A.O.

; APPLICANT: Eldhinn, Deirdre Ni

; APPLICANT: Perkins, Samuel L.

; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus

; FILE REFERENCE: P06283US2/BAS

; CURRENT APPLICATION NUMBER: US/09/200,650E

; CURRENT FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: 60/066,815

; PRIOR FILING DATE: 1997-11-26

; PRIOR APPLICATION NUMBER: 60/098,427

; PRIOR FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 1166

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-200-650E-7

Query Match 7.0%; Score 157.5; DB 5; Length 1166;

Best Local Similarity 21.6%; Pred. No. 0.0052;

Matches 126; Conservative 74; Mismatches 192; Indels 191; Gaps 32;

Qy 2 VENKEETPETDSEEEVITKANLIFANGSTQTAEE-----FKGTPEKATSE--AYAYAD 54

Db 317 IDNKVKKGMTINYNKVN-IPSDLTQDKNPIDITDPSGEVIAGFDKATKITYTFTD 375

Qy 55 TLKKNDEYTVDVADK-----GY-----TLNKFAGKEKTPPEPKKEVTI--KANL 98

Db 376 YVDKYE-----DIKSLTLYSYDKKTPVNETSILNLTFAFAGK---ETSONVTVDYQDPM 427

Qy 99 IYADGKTQTAEEFGTPEEATAEAYRVADALKKNGEYTVDA-----DKGYTNIRKPAK 153

Db 428 VHGDSNQISFTLDEKQDIEQIYVNPPLKKSATNPKVDIAGSQVDYDYG---NKL-GN 483

Qy 154 EKTPEEPKEEVTI-----KANLIY-----172

Db 484 GSTIIDQNTKIKYKVNSDQQLPQSNRIYDFSOYEDVTSQFDNKKFSNNVATLDFGDN 543

Qy 173 -----ADGKTQTAEEFGTPEEATAEAYRVADLLAKEN-----GKYT 208

Db 544 SAVIIVKVSRYTPTSDGELDIAQ--GT-SMRTTDKYGYNVAGYSNFIIVTSNDTGGGDGT 600

Qy 209 VDVADKGYTL-----NIKFAGKEKTPPEPKKEEVTIKANLIYADGKTOTA-----EF 254

Db 601 VKPEEKLYKIGDYVWDVQDQVGGTDSKPKMANVLVTIYDPGTTKSVRTDANGHYEF 660

Qy 255 KG-----TFABATAEAYRYADLLAKENGYTADLEDGGYTNIRFAGKKVDEKPEPM 307

Db 661 GGLKDGETYVVKFETPTGY---LPTKVNCTDGEKDSNGSVTV-----KINGKODMSL 711

Qy 308 DT-----YKLLNGKTLKGETTTEAVDAATAEAKFKQYANNGV--DGEWTVDDATKTF 359

Db 712 DTGYFKPKRYNL---GDYVWEDTNKOGI-----QDANEPGIKDVKVTLKDSTGKV 758

Qy 360 ---TVTEKPEVIDASEL-----TPA---VTTYKLVINGKTLKGETT---KAVDA 400

Db 759 IGTFTTDSKGYKFTLDLNGNYVEFETPAGYTPTVKNTTADKDSNGLTTTGVIKQADN 818

Qy 401 ETAEAKFKQYANNGVDGVWYTD-----DAT-----KTFTVT 432

Db 819 MTLDRGEYKPKYSLGDYVW-YDSNKDGKQDSTEKGIKDVTVT 860

## RESULT 13

US-10-408-765A-1222

; Sequence 1222, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; CURRENT APPLICATION NUMBER: 660088.465

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1222

; LENGTH: 916

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-1222

Query Match 7.0%; Score 156; DB 6; Length 916;

Best Local Similarity 23.5%; Pred. No. 0.005;

Matches 117; Conservative 56; Mismatches 153; Indels 172; Gaps 26;

Qy 1 AVENKEETPETDSEEEVITKANLIFANGSTQTAEEFKGTPEKATSEAYAYADTLKDN 60

Db 491 AAEKEEPEA-----EEEVAAKKSPVKAT-APEVKEEG--EKEEEGQEEEP--BEDE 541

Qy 61 GEYTVADVADGYTLNIRKFKAGKEKTPPEPKKEEVTIKNLIYADGKTQTAEEFKGTPEEATAE 120

Db 542 GAKS-DAEGS-----GSEKESSEKEEGE-----QEEGETE-AEAE--EEAEAK 583

Qy 121 AYRYADALKKNGEYTVDVADKGYTLNIRKFKAGKEKTPPEPKKEEVTIKNLIYADGKTOTA 180

Db 584 -----EKKVKEEKSEEVATKEELV-ADAKVEKP 610

Qy 181 E-PKGTPEEATAEAYRYADLLAKENGYTV---DVADKGYT---LNIRKFKAGKEKTPPEP 232

Db 611 EKAKSPVKSPVE-----EKGKSPVKSPVEEKGKSPVKSPVKSPVEEKGKSPVKSP 660

Qy 233 KEEVTIKNLIYADGKTQTAEE-----KGTEAEATAEAYRYADLLAKENGYTA 281

Db 661 VEE-----KGKSPVKSPVEEKAASPVKSPVEEAKSKA-----EVGKEQ 701

Qy 282 DLEDGGYTNIRFAGK--KYDEKPEEPMDTYKLLINGKTLKGETTTEAVDAATAEAKYF-- 337

Db 702 KEEB---EKEVKEAPKEEVEKKEPKDVPPEKKKAESPVEKEAEVAEAVVTTTKSVKHLE 758

Qy 338 -----KQYANDNG-----VDGEWTVDDATK- 357

Db 759 KETKEEKGKLOQEEKEKEKAGEGEGEGSDKGAGSKRKBEDIAVNGEVEKEEVEQETKE 818

Qy 358 --TFTVTEKPEVIDASELTTPAVTTYKLVINGKTLKGETTTEAVDAATAEAKFAKQYANDNG 415

Db 819 KSGREEEKGVTYVINGLGLSPA-----DEKKGDKSEEEKVVVTKTVEK-----ITSEG 865

Qy 416 VDGWTVDDATKTFVTTE 433

Db 866 GDGATKY--ITKSVTVTQ 881

## RESULT 14

US-60-490-890-475

; Sequence 475, Application US/60490890

; GENERAL INFORMATION:

; APPLICANT: Li, Martha

; APPLICANT: Rupnow, Brent A.

; APPLICANT: Webster, Kevin R.

[illegible]

Search completed: September 3, 2003, 11:14:32  
Job time : 19.6184 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 10:59:36 ; Search time 11.5562 Seconds  
(without alignments)  
1116.704 Million cell updates/sec

Title: US-08-325-278B-1  
Perfect score: 1565  
Sequence: 1 AVENKBPETPTDSEEV.....GGYTINIRPAKKVDEKPEE 305

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1565	100.0	305	2	US-08-795-475-1
2	1565	100.0	434	2	US-08-795-475-3
3	1226.5	78.4	1027	3	US-08-446-137B-2
4	1216	77.7	291	3	US-08-446-137B-4
5	618	39.5	342	3	US-08-828-741B-6
6	618	39.5	342	4	US-09-160-567-6
7	618	39.5	342	4	US-09-710-299-6
8	615	39.3	178	3	US-08-828-741B-13
9	615	39.3	178	4	US-09-160-567-13
10	615	39.3	178	4	US-09-710-299-13
11	615	39.3	198	3	US-08-828-741B-8
12	615	39.3	198	4	US-09-160-567-8
13	615	39.3	198	4	US-09-710-299-8
14	614	39.2	495	3	US-08-828-741B-4
15	614	39.2	495	4	US-09-160-567-4
16	614	39.2	495	4	US-09-710-299-4
17	613	39.2	182	3	US-08-828-741B-2
18	613	39.2	182	4	US-09-160-567-2
19	613	39.2	182	4	US-09-710-299-2
20	334	21.3	75	3	US-08-446-137B-8
21	315.5	20.2	71	3	US-08-446-137B-6
22	306	19.6	74	3	US-08-446-137B-7
23	303	19.4	71	3	US-08-446-137B-5
24	167	10.7	664	3	US-08-669-408B-2
25	151.5	9.7	1183	2	US-08-447-031A-2
26	141.5	9.0	1638	4	US-09-071-035-258
27	141.5	9.0	1638	4	US-09-071-035-262

28	141.5	9.0	1638	4	US-09-071-035-266	Sequence 266, App
29	131.5	8.4	414	4	US-09-286-981B-16	Sequence 16, Appl
30	130.5	8.3	414	4	US-09-286-981B-10	Sequence 10, Appl
31	129	8.2	558	4	US-09-071-035-268	Sequence 268, App
32	128.5	8.2	564	3	US-09-308-022-6	Sequence 6, Appli
33	127.5	8.1	631	3	US-08-847-065-25	Sequence 25, Appl
34	124.5	8.0	669	4	US-09-071-035-264	Sequence 264, App
35	123.5	7.9	952	4	US-09-107-532A-4706	Sequence 4706, Ap
36	123	7.9	936	5	PCT-US93-05944-2	Sequence 2, Appli
37	123	7.9	1008	4	US-09-308-453-2	Sequence 2, Appli
38	122.5	7.8	471	4	US-09-107-532A-5705	Sequence 5705, Ap
39	121.5	7.8	522	3	US-08-961-083-120	Sequence 120, App
40	121.5	7.8	522	4	US-09-536-784-120	Sequence 120, App
41	121.5	7.8	1040	3	US-08-961-083-118	Sequence 118, App
42	121.5	7.8	1040	4	US-09-536-784-118	Sequence 118, App
43	116.5	7.4	734	4	US-09-328-352-4412	Sequence 4412, Ap
44	116.5	7.4	1912	1	US-08-409-995-4	Sequence 4, Appli
45	116.5	7.4	1912	3	US-08-685-467-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-08-795-475-1  
; Sequence 1, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bivrock, Lars  
; APPLICANT: Sjvbring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

Query Match 100.0%; Score 1565; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 5e-126;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AVENKBPETPTDSEEVTKANLIFANGSTQTAEPKGTPEKATSBAYADTLKKDN 60  
|||||

Db 1 AVENKEETPETDSEBEVTIKANLI FANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60  
QY 61 GEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
Db 61 GEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
QY 121 AYRADALKONGEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
Db 121 AYRADALKONGEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
QY 181 EFKGTPEEATAEAYRYADLLAKENGKYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKA 240  
Db 181 EFKGTPEEATAEAYRYADLLAKENGKYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKA 240  
QY 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTADLEDGGYTNIRFAGKKVD 300  
Db 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTADLEDGGYTNIRFAGKKVD 300  
QY 301 EKPEE 305  
Db 301 EKPEE 305

## RESULT 2

US-08-795-475-3  
; Sequence 3, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bjvrck, Lars  
; APPLICANT: Sjvrck, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
US-08-795-475-3

Query Match 100.0%; Score 1565; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 8.1e-126;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEBEVTIKANLI FANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60  
:|||||

Db 1 AVENKEETPETDSEBEVTIKANLI FANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60  
QY 61 GEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
Db 61 GEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
QY 121 AYRADALKONGEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
Db 121 AYRADALKONGEYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
QY 181 EFKGTPEEATAEAYRYADLLAKENGKYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKA 240  
Db 181 EFKGTPEEATAEAYRYADLLAKENGKYTVDVADKGYTLNIFAGKEKTPPEPKBEVTIKA 240  
QY 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTADLEDGGYTNIRFAGKKVD 300  
Db 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTADLEDGGYTNIRFAGKKVD 300  
QY 301 EKPEE 305  
Db 301 EKPEE 305

## RESULT 3

US-08-446-137B-2  
; Sequence 2, Application US/08446137B  
; Patent No. 6162903  
; GENERAL INFORMATION:  
; APPLICANT: Trowern, Angus R.  
; APPLICANT: Atkinson, Anthony  
; APPLICANT: Murphy, Jonathan P.  
; APPLICANT: Laurence, Oliver S.  
; APPLICANT: Dugleby, Clive J.  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; FROM L PROTEIN AND THEIR USES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,137B  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.406  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1027 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-446-137B-2

Query Match 78.4%; Score 1226.5; DB 3; Length 1027;  
Best Local Similarity 79.9%; Pred. No. 2e-96;  
Matches 246; Conservative 22; Mismatches 31; Indels 9; Gaps 4;

QY 2 VENKEETPETDSEBEVTIKANLI FANGSTQTAEPKGTPEKATSEAYAYADTLKKONG 61  
:|||||

Db 276 MERKLSKETPE--PEEVTIKANLIFADGSTQNAEFKGTFAKAVSDAYAYADALKKONG 333  
 QY 62 EYTVADKGYTINIKFAGKKTPEEPKEEVTIKANLIYADGKTQTAFKGTPEEATABA 121  
 Db 334 EYTVADKGLTINIKFAGKKEPEEVTIKVNLIFADGKTQTAFKGTPEEATAKA 393  
 QY 122 YRYADALKKONGEYTVADKGYTINIKFAGKE--KTPPEEKEEVTIKANLIYADGKTQT 179  
 Db 394 YAYADLLAKNGEYTVADLGGTINIKFAGKETPETPEEPEEVTIKVNLIFADGKIQT 453  
 QY 180 AEFKGTPEEATARYADLLAKNGKGYTVADKGYTINIKFAGKE--KTPPEEKEEVT 237  
 Db 454 AEFKGTPEEATAKAYAYANLLAKNGEYTVADLGGTINIKFAGKETPETPEEPEEVT 513  
 QY 238 IKANLIYADGKTQTAFKGTFAEATARYADLLAKNGKGYTVADLGGYTTINIRFAGK 297  
 Db 514 IKVNLIFADGKTQTAFKGTPEEATARYADLLAKNGEYTVADLGGYTTINIKFAGK 573  
 QY 298 KVDEKPEE 305  
 Db 574 ---EQPGE 578

## RESULT 4

US-08-446-137B-4  
 ; Sequence 4, Application US/08446137B  
 ; Patent No. 6162903  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Trowern, Angus R.  
 ; APPLICANT: Ackinson, Anthony  
 ; APPLICANT: Murphy, Jonathan P.  
 ; APPLICANT: Laurence, Oliver S.  
 ; APPLICANT: Duggleby, Clive J.  
 ; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
 ; FROM L PROTEIN AND THEIR USES  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED AND BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/446.137B  
 ; FILING DATE: 22-MAY-1995  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McMasters, David D.  
 ; REGISTRATION NUMBER: 33,963  
 ; REFERENCE/DOCKET NUMBER: 100084.406  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 291 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-446-137B-4

Query Match 77.7%; Score 1216; DB 3; Length 291;  
 Best Local Similarity 82.5%; Pred. No. 2.8e-96;  
 Matches 241; Conservative 20; Mismatches 25; Indels 6; Gaps 3;  
 QY 10 ETPETDSEEVTIKANLIFANGSTQTAFKGTPEEATSEAYADTLKONGEYTVAD 69  
 |||||

Db 2 ETPE--PEEVTIKANLIFADGSTQNAEFKGTFAKAVSDAYAYADALKKONGEYTVAD 59  
 QY 70 KGYTINIKFAGKKTPEEPKEEVTIKANLIYADGKTQTAFKGTPEEATABAYADALK 129  
 Db 60 KGLTINIKFAGKKEPEEVTIKVNLIFADGKTQTAFKGTPEEATAKAYADLLA 119  
 QY 130 KONGEYTVADKGYTINIKFAGKE--KTPPEEKEEVTIKANLIYADGKTQTAFKGTPE 187  
 Db 120 KENGEYTVADLGGTINIKFAGKETPETPEEPEEVTIKVNLIFADGKIQTAFKGTPE 179  
 QY 188 EATAEYRYADLLAKNGKGYTVADKGYTINIKFAGKE--KTPPEEKEEVTIKANLIYA 245  
 Db 180 EATAKAYAYANLLAKNGEYTVADLGGTINIKFAGKETPETPEEPEEVTIKVNLIFA 239  
 QY 246 DGKTQTAFKGTFAEATARYADLLAKNGKGYTVADLGGYTTINIRFAGK 297  
 Db 240 DGKTQTAFKGTPEEATARYADLLAKNGEYTVADLGGYTTINIKFAGK 291

## RESULT 5

US-08-828-741B-6  
 ; Sequence 6, Application US/08828741B  
 ; Patent No. 6043069  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koentgen, Frank  
 ; APPLICANT: Sues, Gabriele M.  
 ; APPLICANT: Tarlinton, David M.  
 ; APPLICANT: Treutlein, Herbert R.  
 ; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
 ; PRODUCING SAME  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
 ; STREET: 400 Garden City Plaza  
 ; CITY: Garden City  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 11530  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/828.741B  
 ; FILING DATE: 26-MAR-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Digiglio, Frank S.  
 ; REGISTRATION NUMBER: 31,346  
 ; REFERENCE/DOCKET NUMBER: 10591  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 742-4343  
 ; TELEFAX: (516) 742-4366  
 ; TELEX: 230 901 SANS UR  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 342 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-828-741B-6

Query Match 39.5%; Score 618; DB 3; Length 342;  
 Best Local Similarity 75.9%; Pred. No. 3.6e-45;  
 Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;  
 QY 16 SEEVTIKANLIFANGSTQTAFKGTPEEATSEAYADTLKONGEYTVADKGYTLN 75  
 Db 174 SAEVTIKANLIFANGSTQTAFKGTPEEATSEAYADTLKONGEYTVADKGYTLN 233  
 QY 76 IKFAGKKTPE-----PKEEVTIKANLIYADGKTQTAFKGTPE 115  
 |||||

Db 234 IKFAGKEATNRNTDSTDYGIQLNSRWGGLTSAEVTIKANLIFANGSTQTAEPKGTPE 293  
 QY 116 EATAEAYRADALKKONGEYTVVDVADKGYTLNKKFAGKEKTPPEPK 161  
 Db 294 KATSEAYAYADTLKKNGBYTVVDVADKGYTLNKKFAGKESAWRHPQ 339

## RESULT 6

US-09-160-567-6  
 ; Sequence 6, Application US/09160567  
 ; Patent No. 6326179  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koentgen, Frank  
 ; APPLICANT: Suess, Gabriele M.  
 ; APPLICANT: Tarlinton, David M.  
 ; APPLICANT: Treutlein, Herbert R.  
 ; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
 ; STREET: 400 Garden City Plaza  
 ; CITY: Garden City  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 11530  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09160,567  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/828,741  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Digiglio, Frank S.  
 ; REGISTRATION NUMBER: 31,346  
 ; REFERENCE/DOCKET NUMBER: 10591  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 742-4343  
 ; TELEFAX: (516) 742-4366  
 ; TELEX: 230 901 SANS UR  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 342 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-160-567-6

Query Match 39.5%; Score 618; DB 4; Length 342;  
 Best Local Similarity 75.9%; Pred. No. 3.6e-45;  
 Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;

QY 16 SEEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKNGBYTVVDVADKGYTLN 75  
 Db 174 SAEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKNGBYTVVDVADKGYTLN 233  
 QY 76 IKFAGKEKTPPE-----PKEEVTIKANLIYADKGTQTAEPKGTPE 115  
 Db 234 IKFAGKEATNRNTDSTDYGIQLNSRWGGLTSAEVTIKANLIFANGSTQTAEPKGTPE 293  
 QY 116 EATAEAYRADALKKONGEYTVVDVADKGYTLNKKFAGKEKTPPEPK 161  
 Db 294 KATSEAYAYADTLKKNGBYTVVDVADKGYTLNKKFAGKESAWRHPQ 339

## RESULT 7

US-09-710-299-6

; Sequence 6, Application US/09710299  
 ; Patent No. 6521741  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koentgen, Frank  
 ; APPLICANT: Suess, Gabriele M.  
 ; APPLICANT: Tarlinton, David M.  
 ; APPLICANT: Treutlein, Herbert R.  
 ; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
 ; STREET: 400 Garden City Plaza  
 ; CITY: Garden City  
 ; STATE: New York  
 ; COUNTRY: United States of America  
 ; ZIP: 11530  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09710,299  
 ; FILING DATE: 09-No. 6521741-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/828,741  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Digiglio, Frank S.  
 ; REGISTRATION NUMBER: 31,346  
 ; REFERENCE/DOCKET NUMBER: 10591  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 742-4343  
 ; TELEFAX: (516) 742-4366  
 ; TELEX: 230 901 SANS UR  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 342 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 ; US-09-710-299-6

Query Match 39.5%; Score 618; DB 4; Length 342;  
 Best Local Similarity 75.9%; Pred. No. 3.6e-45;  
 Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;

QY 16 SEEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKNGBYTVVDVADKGYTLN 75  
 Db 174 SAEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKNGBYTVVDVADKGYTLN 233  
 QY 76 IKFAGKEKTPPE-----PKEEVTIKANLIYADKGTQTAEPKGTPE 115  
 Db 234 IKFAGKEATNRNTDSTDYGIQLNSRWGGLTSAEVTIKANLIFANGSTQTAEPKGTPE 293  
 QY 116 EATAEAYRADALKKONGEYTVVDVADKGYTLNKKFAGKEKTPPEPK 161  
 Db 294 KATSEAYAYADTLKKNGBYTVVDVADKGYTLNKKFAGKESAWRHPQ 339

## RESULT 8

US-08-828-741B-13  
 ; Sequence 13, Application US/08828741B  
 ; Patent No. 6043069  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koentgen, Frank  
 ; APPLICANT: Suess, Gabriele M.  
 ; APPLICANT: Tarlinton, David M.  
 ; APPLICANT: Treutlein, Herbert R.  
 ; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF



```

; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-828-741B-13

Query Match 39.3%; Score 615; DB 3; Length 178;
Best Local Similarity 76.2%; Pred. No. 2.7e-45;
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 BEVTIKANLI FANGSTQTAEFGTPEKATSEAYAYADTLKKONGEYTVDVADKGYTLNIK 77
DB 12 BEVTIKANLI FANGSTQTAEFGTPEKATSEAYAYADTLKKONGEYTVDVADKGYTLNIK 71

QY 78 FAGKEKTPPEE-----PKBEVTIKANLIYADGKTQTAEFGKTFPEA 117
DB 72 FAGKEATNRNTDGTGYILQINRWGGLTSAEVTIKANLI FANGSTQTAEFGKTFPEKA 131

QY 118 TAEAYRYADALKKONGEYTVDVADKGYTLNIK FAKGKTPPEPK 161
DB 132 TSEAYAYADTLKKONGEYTVDVADKGYTLNIK FAKGESAWRHPQ 175

RESULT 9
US-09-160-567-13
; Sequence 13, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/710,299
; FILING DATE: 09-NOV-2000
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE: <unknown>

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-160-567-13

Query Match 39.3%; Score 615; DB 4; Length 178;
Best Local Similarity 76.2%; Pred. No. 2.7e-45;
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 BEVTIKANLI FANGSTQTAEFGTPEKATSEAYAYADTLKKONGEYTVDVADKGYTLNIK 77
DB 12 BEVTIKANLI FANGSTQTAEFGTPEKATSEAYAYADTLKKONGEYTVDVADKGYTLNIK 71

QY 78 FAGKEKTPPEE-----PKBEVTIKANLIYADGKTQTAEFGKTFPEA 117
DB 72 FAGKEATNRNTDGTGYILQINRWGGLTSAEVTIKANLI FANGSTQTAEFGKTFPEKA 131

QY 118 TAEAYRYADALKKONGEYTVDVADKGYTLNIK FAKGKTPPEPK 161
DB 132 TSEAYAYADTLKKONGEYTVDVADKGYTLNIK FAKGESAWRHPQ 175

RESULT 10
US-09-710-299-13
; Sequence 13, Application US/09710299
; Patent No. 6521741
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/710,299
; FILING DATE: 09-NOV-2000
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE: <unknown>

```

ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 178 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-710-299-13

Query Match 39.3%; Score 615; DB 4; Length 178;  
Best Local Similarity 76.2%; Pred. No. 2.7e-45;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 EVVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLNIK 77  
Db 12 EVVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLNIK 71  
QY 78 FAGKEKTPEE-----PKEEVTIKANLIYADGKTQTAEFKGTPEEA 117  
Db 72 FAGKEATNRNTDGTGYILQINRWGGLTSAEVTIKANLIFANGSTQTAEPKGTPEKA 131  
QY 118 TAEAYRYADALKKONGEYTVDVADKGYTLNIKFKAGKEKTPPEPK 161  
Db 132 TSEAYAYADTLKKDNGEYTVDVADKGYTLNIKFKAGKESAWRHPQ 175

RESULT 11  
US-08-828-741B-8  
; Sequence 8, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,741B  
; FILING DATE: 26-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-828-741B-8

Query Match 39.3%; Score 615; DB 3; Length 198;  
Best Local Similarity 76.2%; Pred. No. 3.1e-45;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 EVVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLNIK 77  
Db 32 EVVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDNGEYTVDVADKGYTLNIK 91  
QY 78 FAGKEKTPEE-----PKEEVTIKANLIYADGKTQTAEFKGTPEEA 117  
Db 92 FAGKEATNRNTDGTGYILQINRWGGLTSAEVTIKANLIFANGSTQTAEPKGTPEKA 151  
QY 118 TAEAYRYADALKKONGEYTVDVADKGYTLNIKFKAGKEKTPPEPK 161  
Db 152 TSEAYAYADTLKKDNGEYTVDVADKGYTLNIKFKAGKESAWRHPQ 195

RESULT 12  
US-09-160-567-8  
; Sequence 8, Application US/09160567  
; Patent No. 6326179  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/160,567  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-160-567-8

Query Match 39.3%; Score 615; DB 4; Length 198;  
Best Local Similarity 76.2%; Pred. No. 3.1e-45;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 EVVTIKANLIPANGSTQTAEPFGTPEKATSEAYAYADTLKONGEYTVDVADKGYTLNFK 77  
DB 32 EVVTIKANLIPANGSTQTAEPFGTPEKATSEAYAYADTLKONGEYTVDVADKGYTLNFK 91  
QY 78 FAGKEKTPPEE-----PKEEVTIKANLIYADGKTOTABFKGTPEEA 117  
DB 92 FAGKEATNRTDGTGYILOINRWGGLTSABEVTIKANLIPANGSTQTAEPFGTPEKA 151  
QY 118 TABAYRYADALKKONGEYTVDVADKGYTLNFKPAGKEKTPPEPK 161  
DB 152 TSEAYAYADTLKONGEYTVDVADKGYTLNFKPAGKESAWRHPQ 195  
RESULT 13  
US-09-710-299-8  
; Sequence 8, Application US/09710299  
; Patent No. 6521741  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; Suesse, Gabriele M.  
; Tarlington, David M.  
; Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/710,299  
; FILING DATE: 09-NO. 6521741-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-710-299-8  
Query Match 39.3%; Score 615; DB 4; Length 198;  
Best Local Similarity 76.2%; Pred. No. 3.1e-45;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;  
QY 18 EVVTIKANLIPANGSTQTAEPFGTPEKATSEAYAYADTLKONGEYTVDVADKGYTLNFK 77  
DB 32 EVVTIKANLIPANGSTQTAEPFGTPEKATSEAYAYADTLKONGEYTVDVADKGYTLNFK 91  
QY 78 FAGKEKTPPEE-----PKEEVTIKANLIYADGKTOTABFKGTPEEA 117  
DB 92 FAGKEATNRTDGTGYILOINRWGGLTSABEVTIKANLIPANGSTQTAEPFGTPEKA 151

QY 118 TABAYRYADALKKONGEYTVDVADKGYTLNFKPAGKEKTPPEPK 161  
DB 152 TSEAYAYADTLKONGEYTVDVADKGYTLNFKPAGKESAWRHPQ 195  
RESULT 14  
US-08-828-741B-4  
; Sequence 4, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; Suesse, Gabriele M.  
; Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,741B  
; FILING DATE: 26-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 495 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-828-741B-4  
Query Match 39.2%; Score 614; DB 3; Length 495;  
Best Local Similarity 78.6%; Pred. No. 1.3e-44;  
Matches 125; Conservative 4; Mismatches 10; Indels 20; Gaps 1;  
QY 16 SEEVTIKANLIPANGSTQTAEPFGTPEKATSEAYAYADTLKONGEYTVDVADKGYTLN 75  
DB 174 SAEVTIKANLIPANGSTQTAEPFGTPEKATSEAYAYADTLKONGEYTVDVADKGYTLN 233  
QY 76 IKFAGKEKTPPEE-----PKEEVTIKANLIYADGKTOTABFKGTPE 115  
DB 234 IKFAGKEATNRTDGTGYILOINRWGGLTSABEVTIKANLIPANGSTQTAEPFGTPE 293  
RESULT 15  
US-09-160-567-4  
; Sequence 4, Application US/09160567  
; Patent No. 6326179  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; Suesse, Gabriele M.

APPLICANT: Tarlinton, David M.  
APPLICANT: Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
TITLE OF INVENTION: PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 495 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-4

Query Match 39.2%; Score 614; DB 4; Length 495;  
Best Local Similarity 78.6%; Pred. NO. 1.3e-44;  
Matches 125; Conservative 4; Mismatches 10; Indels 20; Gaps 1;  
QY 16 SEEEVTKANLIFANGSTQTAEPKGTSEAYAYADTLKKNNGEYTVDVADKGYTLN 75  
Db 174 SAEVTKANLIFANGSTQTAEPKGTSEAYAYADTLKKNNGEYTVDVADKGYTLN 233  
QY 76 IKFAGKEKTPPE-----PKEVTKANLIYADGKTQTAEPKGTPE 115  
Db 234 IKFAGKEATNRNTDGSYDYLQINRWGLTSAEVTKANLIFANGSTQTAEPKGTPE 293  
QY 116 EATAQAYRYADALKKNGEYTVDVADKGYTLNKFAGKE 154  
Db 294 KATSEAYAYADTLKKNNGEYTVDVADKGYTLNKFAGKE 332

Search completed: September 3, 2003, 11:04:45  
Job time: 12.5562 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:03:32 ; Search time 31.7794 Seconds  
(without alignments)  
1317.758 Million cell updates/sec

Title: US-08-325-278B-1

Perfect score: 1565

Sequence: 1 AVENKBPETPTDSEEV.....GGYTINIRFAGKKVDEKPEE 305

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgm2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgm2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgm2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgm2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgm2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgm2\_6/ptodata/1/pubpaa/PCT05\_PUBCOMB.pep.\*
- 7: /cgm2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgm2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgm2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgm2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgm2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgm2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgm2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgm2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgm2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgm2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgm2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgm2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1565	100.0	305	US-08-325-278-1	Sequence 1, Appli
2	1565	100.0	434	US-08-325-278-3	Sequence 3, Appli
3	618	39.5	342	US-10-345-618-6	Sequence 6, Appli
4	615	39.3	178	US-10-345-618-13	Sequence 13, Appli
5	615	39.3	198	US-10-345-618-8	Sequence 8, Appli
6	614	39.2	495	US-10-345-618-4	Sequence 4, Appli
7	613	39.2	182	US-10-345-618-2	Sequence 2, Appli
8	608	38.8	482	US-10-345-618-16	Sequence 16, Appli
9	389	24.9	76	US-09-808-212A-4	Sequence 4, Appli
10	371	23.7	72	US-09-808-212A-6	Sequence 6, Appli
11	370	23.6	72	US-09-808-212A-10	Sequence 10, Appli
12	369	23.6	72	US-09-808-212A-8	Sequence 8, Appli
13	341	21.8	82	US-09-808-212A-2	Sequence 2, Appli
14	334	21.3	75	US-09-808-212A-18	Sequence 18, Appli
15	315.5	20.2	71	US-09-808-212A-14	Sequence 14, Appli

16	306	19.6	74	10	US-09-808-212A-16	Sequence 16, Appli
17	303	19.4	71	10	US-09-808-212A-12	Sequence 12, Appli
18	166	10.6	669	9	US-09-878-756-4	Sequence 4, Appli
19	154	9.8	1741	10	US-09-871-536-68	Sequence 68, Appli
20	151.5	9.7	1183	10	US-09-870-759-45	Sequence 45, Appli
21	151.5	9.7	1183	12	US-09-751-708A-45	Sequence 45, Appli
22	149	9.5	448	12	US-10-342-224-82	Sequence 82, Appli
23	131.5	8.4	414	12	US-10-254-995-16	Sequence 16, Appli
24	131.5	8.4	836	10	US-09-858-525A-10	Sequence 10, Appli
25	131.5	8.4	871	10	US-09-858-525A-2	Sequence 2, Appli
26	130.5	8.3	376	11	US-09-056-019-7	Sequence 7, Appli
27	130.5	8.3	413	11	US-09-056-019-35	Sequence 35, Appli
28	130.5	8.3	414	12	US-10-254-995-10	Sequence 10, Appli
29	130.5	8.3	581	11	US-09-298-523B-56	Sequence 56, Appli
30	130.5	8.3	663	11	US-09-298-523B-58	Sequence 58, Appli
31	130.5	8.3	663	11	US-09-056-019-8	Sequence 8, Appli
32	130.5	8.3	663	12	US-09-969-748C-14	Sequence 14, Appli
33	129	8.2	506	15	US-10-125-692-18	Sequence 18, Appli
34	128.5	8.2	483	11	US-09-298-523B-10	Sequence 10, Appli
35	127	8.1	481	11	US-09-298-523B-6	Sequence 6, Appli
36	124.5	8.0	1436	15	US-10-080-505-13	Sequence 13, Appli
37	123	7.9	444	11	US-09-298-523B-5	Sequence 5, Appli
38	122	7.8	487	15	US-10-156-761-11764	Sequence 11764, A
39	121.5	7.8	522	9	US-09-765-272-120	Sequence 120, App
40	121.5	7.8	1040	9	US-09-765-272-118	Sequence 118, App
41	117	7.5	439	11	US-09-056-019-30	Sequence 30, Appli
42	117	7.5	1621	15	US-10-185-990-10	Sequence 10, Appli
43	117	7.5	1626	15	US-10-185-990-11	Sequence 11, Appli
44	116.5	7.4	892	12	US-10-205-219-102	Sequence 102, App
45	116.5	7.4	2353	10	US-09-797-862-33	Sequence 33, Appli

#### ALIGNMENTS

RESULT 1  
US-08-325-278-1  
; Sequence 1, Application US/08325278  
; Publication No. US20030027283A1  
; GENERAL INFORMATION:  
; APPLICANT: Bivrock, Lars  
; APPLICANT: Sjvdring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278  
; FILING DATE: 26-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 450023.401  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown

MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
US-08-325-278-1

Query Match 100.0%; Score 1565; DB 8; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1.1e-116;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEVITKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKN 60  
DB 1 AVENKEETPETDSEEVITKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKN 60  
QY 61 GEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
DB 61 GEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
QY 121 AYRYADALKKONGEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
DB 121 AYRYADALKKONGEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
QY 181 EFKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKA 240  
DB 181 EFKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKA 240  
QY 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKVD 300  
DB 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKVD 300  
QY 301 EKPEE 305  
DB 301 EKPEE 305

RESULT 2

US-08-325-278-3  
Sequence 3, Application US/08325278  
Publication No. US20030027283A1  
GENERAL INFORMATION:  
APPLICANT: Bjvrck, Lars  
APPLICANT: Sjvbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McWaters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 450023.401  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown

MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
US-08-325-278-3

Query Match 100.0%; Score 1565; DB 8; Length 434;  
Best Local Similarity 100.0%; Pred. No. 1.7e-116;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEVITKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKN 60  
DB 1 AVENKEETPETDSEEVITKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKN 60  
QY 61 GEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
DB 61 GEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
QY 121 AYRYADALKKONGEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
DB 121 AYRYADALKKONGEYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKANLIYADGKTQTA 180  
QY 181 EFKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKA 240  
DB 181 EFKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKEKTPPEPKBEVTIKA 240  
QY 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKVD 300  
DB 241 NLIYADGKTQTAEPKGTPEEATAEAYRYADLLAKENGKYTVADVADKGYTLNIPKAGKVD 300  
QY 301 EKPEE 305  
DB 301 EKPEE 305

RESULT 3

US-10-345-618-6  
Sequence 6, Application US/10345618  
Publication No. US20030148484A1  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
APPLICANT: Suess, Gabriele M.  
APPLICANT: Tarlinton, David M.  
APPLICANT: Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
FILE REFERENCE: 13474  
CURRENT APPLICATION NUMBER: US/10/345,618  
CURRENT FILING DATE: 2003-01-16  
PRIOR APPLICATION NUMBER: US/09/509,031  
PRIOR FILING DATE: 2000-06-09  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 6  
LENGTH: 342  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: TLHL protein  
US-10-345-618-6  
Query Match 39.5%; Score 618; DB 12; Length 342;  
Best Local Similarity 75.9%; Pred. No. 1.7e-41;  
Matches 126; Conservative 5; Mismatches 15; Indels 20; Gaps 1;  
QY 16 SEEEVITKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKNGEYTVADVADKGYTLN 75  
DB 174 SAEVITKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKNGEYTVADVADKGYTLN 233  
QY 76 IKPAGKEKTPPEE-----PKEEVTIKANLIYADGKTQTAEPKGTPE 115  
DB 234 IKPAGKEATNRNTOGSTDYGILOINSRWGGUITSAEVITKANLIFANGSTQTAEPKGTPE 293

QY 116 EATAEAYRYADALKKONGEYTVVDVADKGYTLNKKFAGKEKTPPEPK 161  
:|||||  
DB 294 KATSEAYAYADTLKONGEYTVVDVADKGYTLNKKFAGKESAWRHPQ 339  
:|||||

## RESULT 4

US-10-345-618-13  
; Sequence 13, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; PRIOR FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:LHL-omp  
; OTHER INFORMATION: protein sequence  
US-10-345-618-13

Query Match 39.3%; Score 615; DB 12; Length 178;  
Best Local Similarity 76.2%; Pred. No. 1.3e-41;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 BEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKONGEYTVVDVADKGYTLNKK 77  
:|||||  
DB 12 BEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKONGEYTVVDVADKGYTLNKK 71  
:|||||  
QY 78 FAGKEKTPPE-----PKEEVTIKANLIYADGKTQTAEFKGTPEFA 117  
:|||||  
DB 72 FAGKEATNRNTDGSYDGIQINSRWGGTSAEEVTIKANLIFANGSTQTAEFKGTPEFA 131  
:|||||

QY 118 TABEAYRYADALKKONGEYTVVDVADKGYTLNKKFAGKEKTPPEPK 161  
:|||||  
DB 132 TSEAYAYADTLKONGEYTVVDVADKGYTLNKKFAGKESAWRHPQ 175  
:|||||

## RESULT 5

US-10-345-618-8  
; Sequence 8, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; PRIOR FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:LHL.seq  
; OTHER INFORMATION: protein sequence  
US-10-345-618-8

Query Match 39.3%; Score 615; DB 12; Length 198;  
Best Local Similarity 76.2%; Pred. No. 1.5e-41;  
Matches 125; Conservative 5; Mismatches 14; Indels 20; Gaps 1;

QY 18 BEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKONGEYTVVDVADKGYTLNKK 77  
:|||||  
DB 32 BEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKONGEYTVVDVADKGYTLNKK 91  
:|||||  
QY 78 FAGKEKTPPE-----PKEEVTIKANLIYADGKTQTAEFKGTPEFA 117  
:|||||  
DB 92 FAGKEATNRNTDGSYDGIQINSRWGGTSAEEVTIKANLIFANGSTQTAEFKGTPEFA 151  
:|||||  
QY 118 TABEAYRYADALKKONGEYTVVDVADKGYTLNKKFAGKEKTPPEPK 161  
:|||||  
DB 152 TSEAYAYADTLKONGEYTVVDVADKGYTLNKKFAGKESAWRHPQ 195  
:|||||

## RESULT 6

US-10-345-618-4  
; Sequence 4, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; CURRENT FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 495  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:CATAB-TEV  
; OTHER INFORMATION: protein sequence  
US-10-345-618-4

Query Match 39.2%; Score 614; DB 12; Length 495;  
Best Local Similarity 78.6%; Pred. No. 5.8e-41;  
Matches 125; Conservative 4; Mismatches 10; Indels 20; Gaps 1;

QY 16 SEEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKONGEYTVVDVADKGYTLN 75  
:|||||  
DB 174 SAEVTIKANLIFANGSTQTAEFKGTPEKATSEAYAYADTLKONGEYTVVDVADKGYTLN 233  
:|||||  
QY 76 IKFAGKEKTPPE-----PKEEVTIKANLIYADGKTQTAEFKGTPE 115  
:|||||  
DB 234 IKFAGKEATNRNTDGSYDGIQINSRWGGTSAEEVTIKANLIFANGSTQTAEFKGTPE 293  
:|||||  
QY 116 EATAEAYRYADALKKONGEYTVVDVADKGYTLNKKFAGKE 154  
:|||||  
DB 294 KATSEAYAYADTLKONGEYTVVDVADKGYTLNKKFAGKE 332  
:|||||

## RESULT 7

US-10-345-618-2  
; Sequence 2, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618

```
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:LHL protein
; OTHER INFORMATION: sequence
US-10-345-618-2

Query Match          39.2%; Score 613; DB 12; Length 182;
Best Local Similarity 78.5%; Pred. No. 1.9e-41;
Matches 124; Conservative 5; Mismatches 9; Indels 20; Gaps 1;

QY 15 DSEEEVTKANLIFANGSTQTAFFKGTFFKATSEAYAYADTLKKONGEYTVDVADKGYTL 74
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 25 DNTVEVTKANLIFANGSTQTAFFKGTFFKATSEAYAYADTLKKONGEYTVDVADKGYTL 84
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 75 NIKPAGKEKTPREP-----KKEVTKANLIYADGKTQTAFFKGTFF 114
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 85 NIKFAGKEATNRNTDGSTDYGILQINRWGGLTKKEVTKANLIFANGSTQTAFFKGTFF 144
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 115 EATAEAYADALKKONGEYTVDVADKGYTLNPKFAG 152
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 145 EKATSEAYADTLKKONGEYTVDVADKGYTLNPKFAG 182
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
US-10-345-618-16
; Sequence 16, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suesse, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:comTLgL protein
; OTHER INFORMATION: sequence
US-10-345-618-16

Query Match          38.8%; Score 608; DB 12; Length 482;
Best Local Similarity 80.0%; Pred. No. 1.7e-40;
Matches 124; Conservative 5; Mismatches 10; Indels 16; Gaps 1;

QY 16 SBEVEVTKANLIFANGSTQTAFFKGTFFKATSEAYAYADTLKKONGEYTVDVADKGYTLN 75
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 321 SAEVEVTKANLIFANGSTQTAFFKGTFFKATSEAYAYADTLKKONGEYTVDVADKGYTLN 380
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 76 IKPAGKEK-----PEEPKEVTKANLIYADGKTQTAFFKGTFF 119
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 381 IKFAGKEASGGGGGGGGGGGGGSAEEVTKANLIFANGSTQTAFFKGTFF 440
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 120 EAYVADALKKONGEYTVDVADKGYTLNPKFAGKE 154
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 441 EAYVADTLKKONGEYTVDVADKGYTLNPKFAGKE 475
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
RESULT 9
US-09-808-212A-4
; Sequence 4, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-4

Query Match          24.9%; Score 389; DB 10; Length 76;
Best Local Similarity 100.0%; Pred. No. 3.6e-24;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KEETPETPTDSEEEVTKANLIFANGSTQTAFFKGTFFKATSEAYAYADTLKKONGEYTV 64
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 KEETPETPTDSEEEVTKANLIFANGSTQTAFFKGTFFKATSEAYAYADTLKKONGEYTV 60
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 65 VDVADKGYTLNPKFAG 80
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 VDVADKGYTLNPKFAG 76
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
US-09-808-212A-6
; Sequence 6, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-6

Query Match          23.7%; Score 371; DB 10; Length 72;
Best Local Similarity 100.0%; Pred. No. 9.1e-23;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 KEKTPPEEPKEEVEVTKANLIYADGKTQTAFFKGTFFKATSEAYAYADALKKONGEYTV 140
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 KEKTPPEEPKEEVEVTKANLIYADGKTQTAFFKGTFFKATSEAYAYADALKKONGEYTV 60
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 141 DKGYTLNPKFAG 152
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 DKGYTLNPKFAG 72
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 11
US-09-808-212A-10
; Sequence 10, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
```



```

; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-10

Query Match      23.6%; Score 370; DB 10; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KETPEEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDABLE 284
DB 1 KETPEEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDABLE 60

QY 285 DGGYTINIRFAG 296
DB 61 DGGYTINIRFAG 72

RESULT 12
US-09-808-212A-8
; Sequence 8, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-8

Query Match      23.6%; Score 369; DB 10; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 KETPEEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDVA 212
DB 1 KETPEEPKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDVA 60

QY 213 DKGYTLNKFAG 224
DB 61 DKGYTLNKFAG 72

RESULT 13
US-09-808-212A-2
; Sequence 2, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-2

Query Match      21.8%; Score 341; DB 10; Length 82;
Best Local Similarity 82.9%; Pred. No. 2.6e-20;
Matches 68; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 218 LNIKFAGKE--KTPPEPKKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKE 275
DB 1 MNIFPAGKETPTPEPKKEEVTIKVNLIFADGKTQTAEFKGTFAEATAEAYRYADLLAKV 60

QY 276 NGKYTDABLEDGGYTINIRFAGK 297
DB 61 NGEYTDABLEDGGYHNINIFAGK 82

RESULT 14
US-09-808-212A-18
; Sequence 18, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-18

Query Match      21.3%; Score 334; DB 10; Length 75;
Best Local Similarity 90.1%; Pred. No. 8.3e-20;
Matches 64; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 227 KTPPEPKKEEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENGKYTDLEDG 286
DB 5 ETPEPKKEEVTIKVNLIFADGKTQTAEFKGTFAEATAEAYRYADLLAKVNGEYTDLEDG 64

QY 287 GYTINIRFAGK 297
DB 65 GYTINIRFAGK 75

RESULT 15
US-09-808-212A-14
; Sequence 14, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-14

```

Query Match 20.2%; Score 315.5; DB 10; Length 71;  
Best Local Similarity 87.5%; Pred. No. 2.3e-18;  
Matches 63; Conservative 4; Mismatches 4; Indels 1; Gaps 1;  
QY 225 KEKTPPEPKPEVTIKANLIYADGKTQTAEFKGTFABATAAYRYADLLAKENGKYTDLE 284  
Db 1 KEK-PEEPKPEVTIKVNLIFADGKTQTAEFKGTFEATAKAYAYADLLAKENGKYTDLE 59  
QY 285 DGGYTINIRFAG 296  
Db 60 DGGTINIRKIFAG 71

Search completed: September 3, 2003, 11:15:53  
Job time : 32.7794 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:01:51 ; Search time 222.869 Seconds  
(without alignments)  
1191.101 Million cell updates/sec

Title: US-08-325-278B-1

Perfect score: 1565

Sequence: 1 AVENKBPETPTDSREEV.....GGYINIRPAGKVKDEKPEE 305

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS COMB.ppe:\*  
2: /cgn2\_6/ptodata/1/paa/US06 COMB.ppe:\*  
3: /cgn2\_6/ptodata/1/paa/US07 COMB.ppe:\*  
4: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
5: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
6: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
7: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
8: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
9: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
10: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
11: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
12: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
13: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
14: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
15: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
16: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
17: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
18: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
19: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
20: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
21: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
22: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
23: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
24: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
25: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
26: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
27: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
28: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
29: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
30: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*  
31: /cgn2\_6/ptodata/1/paa/US08 COMB.ppe:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1565	100.0	305	7	US-08-325-278-1 Sequence 1, Appli
2	1565	100.0	305	7	US-08-325-278A-1 Sequence 1, Appli

3	1565	100.0	305	7	US-08-325-278B-1 Sequence 1, Appli
4	1565	100.0	434	7	US-08-325-278-3 Sequence 3, Appli
5	1565	100.0	434	7	US-08-325-278A-3 Sequence 3, Appli
6	1565	100.0	434	7	US-08-325-278B-3 Sequence 3, Appli
7	1561	99.7	719	22	US-09-791-537-10210 Sequence 10210, A
8	1561	99.7	719	22	US-09-791-537-96101 Sequence 96101, A
9	1544.5	98.7	467	25	US-09-980-469-12 Sequence 12, Appli
10	1366	87.3	367	23	US-09-889-182A-4 Sequence 4, Appli
11	1226.5	78.4	992	22	US-09-791-537-88366 Sequence 88366, A
12	1226.5	78.4	1027	7	US-08-331-637-2 Sequence 2, Appli
13	1226.5	78.4	1027	8	US-08-446-137A-2 Sequence 2, Appli
14	1226.5	78.4	1027	15	US-09-187-295-2 Sequence 2, Appli
15	1216	77.7	291	8	US-08-446-137A-4 Sequence 4, Appli
16	618	39.5	342	19	US-09-509-031-6 Sequence 6, Appli
17	618	39.5	342	23	US-09-820-048A-6 Sequence 6, Appli
18	615	39.3	178	19	US-09-509-031-13 Sequence 13, Appli
19	615	39.3	178	23	US-09-820-048A-13 Sequence 13, Appli
20	615	39.3	198	19	US-09-509-031-8 Sequence 8, Appli
21	615	39.3	198	23	US-09-820-048A-8 Sequence 8, Appli
22	614	39.2	495	19	US-09-509-031-4 Sequence 4, Appli
23	614	39.2	495	23	US-09-820-048A-4 Sequence 4, Appli
24	613	39.2	182	19	US-09-509-031-2 Sequence 2, Appli
25	613	39.2	182	23	US-09-820-048A-2 Sequence 2, Appli
26	608	38.8	482	19	US-09-509-031-16 Sequence 16, Appli
27	400	25.6	78	22	US-09-791-537-22553 Sequence 22553, A
28	389	24.9	76	23	US-09-808-212A-6 Sequence 6, Appli
29	371	23.7	72	23	US-09-808-212A-4 Sequence 4, Appli
30	370	23.6	72	23	US-09-808-212A-10 Sequence 10, Appli
31	369	23.6	72	23	US-09-808-212A-2 Sequence 2, Appli
32	341	21.8	82	23	US-09-808-212A-18 Sequence 18, Appli
33	334	21.3	75	23	US-09-808-212A-14 Sequence 14, Appli
34	315.5	20.2	71	23	US-09-808-212A-16 Sequence 16, Appli
35	306	19.6	74	23	US-09-808-212A-12 Sequence 12, Appli
36	303	19.4	71	23	US-09-808-212A-12 Sequence 12, Appli
37	167	10.7	664	10	US-08-669-408-2 Sequence 2, Appli
38	167	10.7	664	23	US-09-878-756-4 Sequence 4, Appli
39	166	10.6	669	23	US-09-878-756-4 Sequence 4, Appli
40	156	10.0	1728	28	US-10-282-122A-56997 Sequence 56997, A
41	156	10.0	1747	15	US-09-134-000-5999 Sequence 5999, Ap
42	156	10.0	1747	15	US-09-134-000C-5999 Sequence 5999, Ap
43	156	10.0	1747	30	US-10-434-665-5999 Sequence 2, Appli
44	154.5	9.9	1156	3	US-07-861-804-2 Sequence 2, Appli
45	154.5	9.9	1156	8	US-08-447-031-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-325-278-1  
; Sequence 1, Application US/08325278  
; GENERAL INFORMATION:  
; APPLICANT: Bjvck, Lars  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278  
; FILING DATE: 26-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278A-1

Query Match      100.0%; Score 1565; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 4.7e-125;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETPTDSEEEVTTKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60
DB 1 AVENKEETPETPTDSEEEVTTKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60
QY 61 GEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120
DB 61 GEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120
QY 121 AYRYADALKKONGEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180
DB 121 AYRYADALKKONGEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180
QY 181 EFKGTFPEATAEAYRYADLLAKENGKYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240
DB 181 EFKGTFPEATAEAYRYADLLAKENGKYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240
QY 241 NLIYADGKTQTAEPKGTPEATAEAYRYADLLAKENGKYTADLEGGYTNIRFAGKKVD 300
DB 241 NLIYADGKTQTAEPKGTPEATAEAYRYADLLAKENGKYTADLEGGYTNIRFAGKKVD 300
QY 301 EKPEE 305
DB 301 EKPEE 305

RESULT 2
US-08-325-278A-1
; Sequence 1, Application US/08325278A
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278A
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.

; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278A-1

Query Match      100.0%; Score 1565; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 4.7e-125;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETPTDSEEEVTTKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60
DB 1 AVENKEETPETPTDSEEEVTTKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKDN 60
QY 61 GEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120
DB 61 GEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120
QY 121 AYRYADALKKONGEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180
DB 121 AYRYADALKKONGEYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180
QY 181 EFKGTFPEATAEAYRYADLLAKENGKYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240
DB 181 EFKGTFPEATAEAYRYADLLAKENGKYTVADVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240
QY 241 NLIYADGKTQTAEPKGTPEATAEAYRYADLLAKENGKYTADLEGGYTNIRFAGKKVD 300
DB 241 NLIYADGKTQTAEPKGTPEATAEAYRYADLLAKENGKYTADLEGGYTNIRFAGKKVD 300
QY 301 EKPEE 305
DB 301 EKPEE 305

RESULT 3
US-08-325-278B-1
; Sequence 1, Application US/08325278B
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278B
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
```

REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 100084.402  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-08-325-278B-1

Query Match 100.0%; Score 1565; DB 7; Length 305;  
Best Local Similarity 100.0%; Pred. No. 4.7e-125;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEBEVTKANLIIFANGSTQTAEPKGTPEKATSBAYAYADTLKKDN 60  
DB 1 AVENKEETPETDSEBEVTKANLIIFANGSTQTAEPKGTPEKATSBAYAYADTLKKDN 60

QY 61 GEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
DB 61 GEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120

QY 121 AYRYADALKONGEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180  
DB 121 AYRYADALKONGEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180

QY 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240  
DB 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240

QY 241 NLIYADGKTQTAEPKGTFAEATAEAYRYADLLAKENGYTADLEDGGYTTINIRFAGKKVD 300  
DB 241 NLIYADGKTQTAEPKGTFAEATAEAYRYADLLAKENGYTADLEDGGYTTINIRFAGKKVD 300

QY 301 EKPEE 305  
DB 301 EKPEE 305

RESULT 4  
US-08-325-278-3  
Sequence 3, Application US/08325278  
GENERAL INFORMATION:  
APPLICANT: Björck, Lars  
Sjöbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESS: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McMasters, David D.

REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 450023.401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7055  
US-08-325-278-3

Query Match 100.0%; Score 1565; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 7.8e-125;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEBEVTKANLIIFANGSTQTAEPKGTPEKATSBAYAYADTLKKDN 60  
DB 1 AVENKEETPETDSEBEVTKANLIIFANGSTQTAEPKGTPEKATSBAYAYADTLKKDN 60

QY 61 GEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120  
DB 61 GEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAE 120

QY 121 AYRYADALKONGEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180  
DB 121 AYRYADALKONGEYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180

QY 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240  
DB 181 EFKGTFEATAEAYRYADLLAKENGYTVDVADKGYTLNPKAGKEKTPPEPKKEVTIKA 240

QY 241 NLIYADGKTQTAEPKGTFAEATAEAYRYADLLAKENGYTADLEDGGYTTINIRFAGKKVD 300  
DB 241 NLIYADGKTQTAEPKGTFAEATAEAYRYADLLAKENGYTADLEDGGYTTINIRFAGKKVD 300

QY 301 EKPEE 305  
DB 301 EKPEE 305

RESULT 5  
US-08-325-278A-3  
Sequence 3, Application US/08325278A  
GENERAL INFORMATION:  
APPLICANT: Björck, Lars  
Sjöbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESS: Seed IP Law Group  
STREET: 701 Fifth Avenue Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278A  
FILING DATE: 26-Oct-1994  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 100084.402  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 434 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-08-325-278A-3

Query Match 100.0%; Score 1565; DB 7; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-125;  
 Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKDN 60  
 DB 1 AVENKEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKDN 60  
 QY 61 GEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTAFFKGTPEEATAE 120  
 DB 61 GEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTAFFKGTPEEATAE 120  
 QY 121 AYRADALKKONGEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180  
 DB 121 AYRADALKKONGEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180  
 QY 181 EFKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKA 240  
 DB 181 EFKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKA 240  
 QY 241 NLIYADGKTQTAFFKGTPEEATAEAYRYADLLAKENGYTADLEDGGYTNIRFAGKKVD 300  
 DB 241 NLIYADGKTQTAFFKGTPEEATAEAYRYADLLAKENGYTADLEDGGYTNIRFAGKKVD 300  
 QY 301 EKPEE 305  
 DB 301 EKPEE 305

#### RESULT 6

US-08-325-278B-3

Sequence 3, Application US/08325278B  
 GENERAL INFORMATION:  
 APPLICANT: Bjorck, Lars  
 Sjoerding, Ulf  
 TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seed IP Law Group  
 STREET: 701 Fifth Avenue Suite 6300  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION NUMBER: US/08/325,278B  
 APPLICATION NUMBER: US/08/325,278B  
 FILING DATE: 26-Oct-1994  
 CLASSIFICATION: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Potter, Jane E. R.  
 REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 100084.402  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 434 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-08-325-278B-3

Query Match 100.0%; Score 1565; DB 7; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-125;  
 Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVENKEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKDN 60  
 DB 1 AVENKEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKDN 60  
 QY 61 GEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTAFFKGTPEEATAE 120  
 DB 61 GEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTAFFKGTPEEATAE 120  
 QY 121 AYRADALKKONGEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180  
 DB 121 AYRADALKKONGEYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKANLIYADGKTQTA 180  
 QY 181 EFKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKA 240  
 DB 181 EFKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNFKPAGKEKTPPEPKKEVTIKA 240  
 QY 241 NLIYADGKTQTAFFKGTPEEATAEAYRYADLLAKENGYTADLEDGGYTNIRFAGKKVD 300  
 DB 241 NLIYADGKTQTAFFKGTPEEATAEAYRYADLLAKENGYTADLEDGGYTNIRFAGKKVD 300  
 QY 301 EKPEE 305  
 DB 301 EKPEE 305

#### RESULT 7

US-09-791-537-10210

Sequence 10210, Application US/09791537  
 GENERAL INFORMATION:  
 APPLICANT: Bionomix, Inc.  
 APPLICANT: Debe, Derek  
 APPLICANT: Danzer, Joseph  
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
 TITLE OF INVENTION: METHODS OF USE THEREOF  
 FILE REFERENCE: 261/210  
 CURRENT APPLICATION NUMBER: US/09/791,537  
 CURRENT FILING DATE: 2001-02-22  
 NUMBER OF SEQ ID NOS: 153055  
 SOFTWARE: Patent in version 3.0  
 SEQ ID NO 10210  
 LENGTH: 719  
 TYPE: PRT  
 ORGANISM: Peptostreptococcus magnus

US-09-791-537-10210

Query Match 99.7%; Score 1561; DB 22; Length 719;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-124;  
 Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VENKEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKONG 61  
 DB 95 VENKEETPETDSEEEVTKANLIFANGSTQTAFFKGTPEKATSEAYAYADTLKKONG 154

QY 62 EYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEA 121  
DB 155 EYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEA 214  
QY 122 YRYADALKDNGEYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAE 181  
DB 215 YRYADALKDNGEYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAE 274  
QY 182 FKGTFEATAEAAYRYADLLAKENGKYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKAN 241  
DB 275 FKGTFEATAEAAYRYADLLAKENGKYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKAN 334  
QY 242 LIYADGKTQTAEFKGTFEATAEAAYRYADLLAKENGKYTVADLEGGYTTINIRFAGKKVDE 301  
DB 335 LIYADGKTQTAEFKGTFEATAEAAYRYADLLAKENGKYTVADLEGGYTTINIRFAGKKVDE 394  
QY 302 KPBE 305  
DB 395 KPBE 398

## RESULT 8

US-09-791-537-96101

; Sequence 96101, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 96101

; LENGTH: 719

; TYPE: PRT

; ORGANISM: Peptostreptococcus magnus

US-09-791-537-96101

Query Match 99.7%; Score 1561; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 3.5e-124;  
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VENKEETPETPTDSEEEVTIKANLI FANGSTQTAEFKGTFREKATSEAYAYADTLKKONG 61  
DB 95 VENKEETPETPTDSEEEVTIKANLI FANGSTQTAEFKGTFREKATSEAYAYADTLKKONG 154  
QY 62 EYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEA 121  
DB 155 EYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEA 214  
QY 122 YRYADALKDNGEYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAE 181  
DB 215 YRYADALKDNGEYTVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAE 274  
QY 182 FKGTFEATAEAAYRYADLLAKENGKYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKAN 241  
DB 275 FKGTFEATAEAAYRYADLLAKENGKYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKAN 334  
QY 242 LIYADGKTQTAEFKGTFEATAEAAYRYADLLAKENGKYTVADLEGGYTTINIRFAGKKVDE 301  
DB 335 LIYADGKTQTAEFKGTFEATAEAAYRYADLLAKENGKYTVADLEGGYTTINIRFAGKKVDE 394  
QY 302 KPBE 305  
DB 395 KPBE 398

## RESULT 9

US-09-980-469-12

Query Match 87.3%; Score 1366; DB 23; Length 367;  
Best Local Similarity 88.6%; Pred. No. 6.2e-108;  
Matches 264; Conservative 15; Mismatches 15; Indels 4; Gaps 1;

; Sequence 12, Application US/09980469

; GENERAL INFORMATION:

; APPLICANT: Shoseyov, Oded

; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND RECO

; FILE REFERENCE: 01/22924

; CURRENT APPLICATION NUMBER: US/09/980,469

; CURRENT FILING DATE: 2000-06-07

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 467

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: pUC19-cell-Protl-1-cexNG-HDEL fusion encoded product

US-09-980-469-12

Query Match 98.7%; Score 1544.5; DB 25; Length 467;  
Best Local Similarity 99.3%; Pred. No. 4.9e-123;  
Matches 303; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AVENKEETPETPTDSEEEVTIKANLI FANGSTQTAEFKGTFREKATSEAYAYADTLKKON 60  
DB 40 AVENKEETPETPTDSEEEVTIKANLI FANGSTQTAEFKGTFREKATSEAYAYADTLKKON 99  
QY 61 GEYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEA 120  
DB 100 GEYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEA 159  
QY 121 AYRYADALKDNGEYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAE 180  
DB 160 AYRYADALKDNGEYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIKANLIYADGKTQTAE 219  
QY 181 EFKGTFEATAEAAYRYADLL- AKENGKYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIK 239  
DB 220 EFKGTFEATAEAAYRYADLLAKENGKYTVADVADKGYTLNFKAGKEKTPPEPKKEVTIK 279  
QY 240 ANLIYADGKTQTAEFKGTFEATAEAAYRYADLLAKENGKYTVADLEGGYTTINIRFAGKKV 299  
DB 280 ANLIYADGKTQTAEFKGTFEATAEAAYRYADLLAKENGKYTVADLEGGYTTINIRFAGKKV 339  
QY 300 DEKPE 304  
DB 340 DEKPE 344

## RESULT 10

US-09-889-182A-4

; Sequence 4, Application US/09889182A

; GENERAL INFORMATION:

; APPLICANT: Breitling, Frank

; TITLE OF INVENTION: SELECTION OF MONOCLONAL ANTIBODIES

; FILE REFERENCE: 4121-126

; CURRENT APPLICATION NUMBER: US/09/889,182A

; CURRENT FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: pct/ds00/00079

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 367

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Sequence

US-09-889-182A-4





```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1027 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-137A-2

Query Match      78.4%; Score 1226.5; DB 8; Length 1027;
Best Local Similarity 79.9%; Pred. No. 2.3e-95;
Matches 246; Conservative 22; Mismatches 31; Indels 9; Gaps 4;

QY 2 VENKEETPTPTDSEEEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKONG 61
DB 276 MERKLSKEKTPR--PEEEVTIKANLIFADGSTQNAEFKGTFAKAVSDAYAYADALKKONG 333
QY 62 EYTVADVADKGYTLNIFAGKGTPEEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAEA 121
DB 334 EYTVADVADKGLTNLIFAGKKEPEEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAEA 393
QY 122 YRYADALKKONGEYTVADVADKGYTLNIFAGKE--KTPEEPKKEVTIKANLIYADGKTQ 179
DB 394 YAYADLLAKENG EYTDLEDGGNTINIKFAGKETPTPEEPKKEVTIKANLIYADGKIQT 453
QY 180 AEFKGTPEEATAEAAYRYADLLAKENGKYTVADVADKGYTLNIFAGKE--KTPEEPKKEV 237
DB 454 AEFKGTPEEATAKAYAYANLLAKENG EYTDLEDGGNTINIKFAGKETPTPEEPKKEV 513
QY 238 IKANLIYADGKTQTAEPKGTPEEATAEAAYRYADLLAKENGKYTDLEDGGYTTINIRFAGK 297
DB 514 IKVNLIFADGKTQTAEPKGTPEEATAEAAYRYADLLAKENG EYTDLEDGGYTTINIRFAGK 573
QY 298 KVDEKPEE 305
DB 574 ---EQPGE 578

RESULT 15
US-08-446-137A-4
; Sequence 4, Application US/08446137A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM PROTEIN L AND THEIR USES
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137A
; FILING DATE: 22-05-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MITCHARD, LEONARD C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 1418-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-137A-4

Query Match      77.7%; Score 1216; DB 8; Length 291;
Best Local Similarity 82.5%; Pred. No. 2.9e-95;
Matches 241; Conservative 20; Mismatches 25; Indels 6; Gaps 3;

QY 10 ETPETDSEEEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKONGEYTVADVAD 69
DB 2 ETPETDSEEEVTIKANLIFANGSTQTAEPKGTPEKATSEAYAYADTLKKONGEYTVADVAD 59
QY 70 KGYTLNIFAGKKEPTPEEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAEAAYRYADALK 129
DB 60 KGLTNLIFAGKKEPEEPKKEVTIKANLIYADGKTQTAEPKGTPEEATAEAAYRYADLLA 119
QY 130 KONGEYTVADVADKGYTLNIFAGKE--KTPEEPKKEVTIKANLIYADGKTQTAEPKGTPE 187
DB 120 KENG EYTDLEDGGNTINIKFAGKETPTPEEPKKEVTIKANLIYADGKIQTAEFKGTPE 179
QY 188 EATAEAAYRYADLLAKENGKYTVADVADKGYTLNIFAGKE--KTPEEPKKEVTIKANLIYA 245
DB 180 EATAKAYAYANLLAKENG EYTDLEDGGNTINIKFAGKETPTPEEPKKEVTIKANLIYA 239
QY 246 DGTQTAEPKGTFAEATAEAAYRYADLLAKENGKYTDLEDGGYTTINIRFAGK 297

```

Db 240 D G K T O T A E F K G T F E E A T A E A Y R Y A D L L A K V N G E Y T A D L E D G G Y T I N I K P A G K 291

Search completed: September 3, 2003, 11:13:53  
Job time : 224.869 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:02:46 ; Search time 12.3816 Seconds  
(without alignments)  
773.734 Million cell updates/sec

Title: US-08-325-278B-1

Perfect score: 1565

Sequence: 1 AVENKEPTPTDSEEV.....GGYTINIRFAGKVKDEKPEE 305

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 148256 seqs, 31410045 residues

Total number of hits satisfying chosen parameters: 148256

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	141	9.0	1849	6	US-10-637-544-2
2	131.5	8.4	1166	5	US-09-200-650E-7
3	129.5	8.3	1742	6	US-10-615-383-4
4	127.5	8.1	701	6	US-10-333-120A-7
5	127.5	8.1	728	6	US-10-467-534-81
6	126	8.1	630	6	US-10-603-113-20275
7	125	8.0	770	5	US-09-897-516A-4453
8	122	7.6	458	6	US-10-467-421-37
9	119.5	7.6	901	6	US-10-603-113-14747
10	116	7.4	916	6	US-10-408-765A-1222
11	116	7.4	916	7	US-60-490-890-475
12	114.5	7.3	930	5	US-09-200-650E-3
13	114.5	7.3	1394	6	US-10-645-655-2
14	113	7.2	1388	6	US-10-408-765A-1139
15	113	7.2	2060	6	US-10-381-596A-2
16	112	7.2	693	6	US-10-333-120A-10
17	111.5	7.1	718	6	US-10-603-108-2753
18	111.5	7.1	2705	1	PCT-US03-21510-97
19	111.5	7.1	2705	6	US-10-408-765A-2039
20	111.5	7.1	2705	6	US-10-089-320B-44
21	111	7.1	637	6	US-10-617-320-3169
22	111	7.1	864	6	US-10-286-897-2171
23	111	7.1	864	6	US-10-258-898A-2171
24	111	7.1	2138	6	US-10-640-833-5274
25	110	7.0	1315	5	US-09-200-650E-5
26	109.5	7.0	523	6	US-10-333-120A-11

27	109.5	7.0	1939	6	US-10-408-765A-2188	Sequence 2188, Ap
28	109	7.0	1275	6	US-10-408-765A-2051	Sequence 2051, Ap
29	109	7.0	1287	7	US-60-490-890-2181	Sequence 2181, Ap
30	108.5	6.9	818	5	US-09-897-516A-5123	Sequence 5123, Ap
31	108.5	6.9	2647	1	PCT-US03-21510-95	Sequence 95, Appl
32	108	6.9	1079	6	US-10-286-897-3540	Sequence 3540, Ap
33	108	6.9	1079	6	US-10-258-898A-3540	Sequence 3540, Ap
34	108	6.9	1465	6	US-10-276-775-1	Sequence 1, Appl
35	108	6.9	2004	7	US-60-490-890-1750	Sequence 1750, Ap
36	108	6.9	2315	6	US-10-603-114-5434	Sequence 5434, Ap
37	107.5	6.9	1277	6	US-10-326-956-665	Sequence 665, App
38	107	6.8	2515	6	US-10-293-244-3914	Sequence 3914, Ap
39	107	6.8	2515	6	US-10-293-244-3915	Sequence 3915, Ap
40	106.5	6.8	633	6	US-10-613-520-1266	Sequence 1266, Ap
41	106.5	6.8	1452	6	US-10-293-244-1489	Sequence 1489, Ap
42	106.5	6.8	1621	6	US-10-293-244-3457	Sequence 3457, Ap
43	106	6.8	1262	6	US-10-617-320-5067	Sequence 5067, Ap
44	105.5	6.7	1938	6	US-10-408-765A-1168	Sequence 1168, Ap
45	105	6.7	346	5	US-09-674-546A-2551	Sequence 2551, Ap

#### ALIGNMENTS

RESULT 1  
US-10-637-544-2  
; Sequence 2, Application US/10637544  
; GENERAL INFORMATION:  
; APPLICANT: Chr. Hansen A/S  
; TITLE OF INVENTION: Peptides with anti-hypertensive properties  
; FILE REFERENCE: P1032US01  
; CURRENT APPLICATION NUMBER: US/10/637,544  
; CURRENT FILING DATE: 2003-08-11  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1849  
; TYPE: PRT  
; ORGANISM: Lactobacillus helveticus  
US-10-637-544-2

Query Match 9.0%; Score 141; DB 6; Length 1849;  
Best Local Similarity 27.7%; Pred. No. 0.021;  
Matches 76; Conservative 32; Mismatches 108; Indels 58; Gaps 15;

QY	73	TLNKFAGKEKTPPEPKKEVTIKANLIYADGKTQTAEFGTFEEATAA--YRVADALKK 130	
DB	1415	TLNL---DSENVVYNKDKFTI-SGTISDDYKFYDLSINGNDVETSWSAVDYHSEGIKK 1470	
QY	131	DNGEYTVDVADGKGYTLNIK-----FAGKEKTPPEPKKE--VTIKANLIYA 173	
DB	1471	-NFKHEVDLKKGKNTFNVKVTDIQGNSSSQALVWVYEPAKTLAEPSVDKLLTKTANLQL 1529	
QY	174	DKGTQTAEFK-----GTFEEATAEAYRVADLLAKENGKYTVDVADGKY-----TL 218	
DB	1530	KATTDSEAKVYVSLDNGKTFNDVPADGFK-----VTENGTVQFKAVDK-YGNESKVKSV 1583	
QY	219	NIKFGAGKEKTPPEPKKEVTIKANL---IYADGKTQ---TAEFGKTFEATAEAYRVADL 271	
DB	1584	EIKGLNKENQSEDEKELAKENLOAKVDAGEKDLDKYTADSKDFNDALKKA---KDV 1640	
QY	272	LAKENGKYTADLEGGYTTINIRFAGKVKDEKPEE 305	
DB	1641	LADKNNA-LADLQDAAKALD--KAEQALTEKPAE 1671	

RESULT 2  
US-09-200-650E-7  
; Sequence 7, Application US/09200650E  
; GENERAL INFORMATION:  
; APPLICANT: Patti, Joseph M.  
; APPLICANT: Hooker, Timothy J.  
; APPLICANT: Hook, Magnus A.O.







Query Match 7.4%; Score 116; DB 6; Length 916;  
Best Local Similarity 25.0%; Pred. No. 0.55; Indels 118; Gaps 19;  
Matches 81; Conservative 34; Mismatches 91; Indels 118; Gaps 19;

QY 1 AVENKEETPETDSEEVTTKANLIFANGSTQTAEPFGTPEKATSEAYAVADTLKDN 60  
DB AAEKEEPEEA-----EEEVAAKSPVKAT-APEVKEEG--EKEEEOQEEB--EED 541  
QY 61 GYTVVDVADKGYTLNIFAGKEKTPPEKKEVTTKANLIYADGKTQTAEPFGTPEATAE 120  
DB GAKS-DOAEEG-----GSEKESSEKEGE-----QERGETE-AEAE--EAEAK 583  
QY 121 AYRYADALKKONGEYTVVDVADKGYTLNIFAGKEKTPPEKKEVTTKANLIYADGKTQTA 180  
DB -----EKKVEKSEEVATKEELV-ADAKVEKP 610  
QY 181 E-FKGTFFEBATAEAYRYADLLAKENGKTYTV---DVADKGYT---LNIFAGKEKTPPEP 232  
DB EKAKSPVKSPVE-----EKGKSPVKSPVEEKGKSPVKSPVEEKGKSPVKSP 660  
QY 233 KEEVTTKANLIYADGKTQTAEP-----KGTFFABATAEAYRYADLLAKENGKTYTA 281  
DB VEE-----KKGKSPVKSPVEEAKSPVKSPVEEAKSKA-----EYKKGQ 701  
QY 282 DLEDGGYTINIRFAGKVDKPEE 305  
DB 702 KEE-----EKEVKEAPKE 715

RESULT 11  
US-60-490-890-475  
; Sequence 475, Application US/60490890  
; GENERAL INFORMATION:  
; APPLICANT: Li, Martha  
; APPLICANT: Rudnow, Brent A.  
; APPLICANT: Webster, Kevin R.  
; APPLICANT: Jackson, Donald  
; APPLICANT: Wong, Tai W.  
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION  
; FILE REFERENCE: D0310 PSP  
; CURRENT APPLICATION NUMBER: US/60/490,890  
; CURRENT FILING DATE: 2003-07-29  
; NUMBER OF SEQ ID NOS: 2779  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 475  
; LENGTH: 916  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-490-890-475

Query Match 7.4%; Score 116; DB 7; Length 916;  
Best Local Similarity 25.0%; Pred. No. 0.55; Indels 118; Gaps 19;  
Matches 81; Conservative 34; Mismatches 91; Indels 118; Gaps 19;

QY 1 AVENKEETPETDSEEVTTKANLIFANGSTQTAEPFGTPEKATSEAYAVADTLKDN 60  
DB AAEKEEPEEA-----EEEVAAKSPVKAT-APEVKEEG--EKEEEOQEEB--EED 541  
QY 61 GYTVVDVADKGYTLNIFAGKEKTPPEKKEVTTKANLIYADGKTQTAEPFGTPEATAE 120  
DB GAKS-DOAEEG-----GSEKESSEKEGE-----QERGETE-AEAE--EAEAK 583  
QY 121 AYRYADALKKONGEYTVVDVADKGYTLNIFAGKEKTPPEKKEVTTKANLIYADGKTQTA 180  
DB -----EKKVEKSEEVATKEELV-ADAKVEKP 610  
QY 181 E-FKGTFFEBATAEAYRYADLLAKENGKTYTV---DVADKGYT---LNIFAGKEKTPPEP 232  
DB EKAKSPVKSPVE-----EKGKSPVKSPVEEKGKSPVKSPVEEKGKSPVKSP 660  
QY 233 KEEVTTKANLIYADGKTQTAEP-----KGTFFABATAEAYRYADLLAKENGKTYTA 281  
DB VEE-----KKGKSPVKSPVEEAKSPVKSPVEEAKSKA-----EYKKGQ 701

QY 282 DLEDGGYTINIRFAGKVDKPEE 305  
DB 702 KEE-----EKEVKEAPKE 715

RESULT 12  
US-09-200-650E-3  
; Sequence 3, Application US/09200650E  
; GENERAL INFORMATION:  
; APPLICANT: Patti, Joseph M.  
; APPLICANT: Foster, Timothy J.  
; APPLICANT: Hook, Magnus A.O.  
; APPLICANT: Eidiham, Deirdre Ni  
; APPLICANT: Perkins, Samuel L.  
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus  
; FILE REFERENCE: P06283US2/BAS  
; CURRENT APPLICATION NUMBER: US/09/200,650E  
; CURRENT FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: 60/066,815  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/098,427  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 930  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-200-650E-3

Query Match 7.3%; Score 114.5; DB 5; Length 930;  
Best Local Similarity 22.1%; Pred. No. 0.71; Indels 137; Gaps 22;  
Matches 77; Conservative 44; Mismatches 91; Indels 137; Gaps 22;

QY 11 TPETDSEEVTTKANLIFAN-GSTQTAEPFGTPEKATSEAY-----AYADTLKONG-- 61  
DB 407 TPDTSKLADVTDQFDVIYSNDNKTATVDMKGG--QTSSNKQYIIQOVAYPDNSSTDNGKI 464  
QY 62 EYTVVDVADKGYTLNIFKA---GKEKTPPEKKEVTTKANL-----IYADGKTQTAEP-- 109  
DB 465 DYTLDTKYKSWNSYNSVNGSSSTANGDQK-----KYNLGDYWNEDTNKQKQDANEKG 519  
QY 110 FKGTPE-----BEATAE---AYRYADALKKONGEYTVVDVA-DKGYTLNIFAGKE 154  
DB 520 IKGVVVILKDSNGKELDRITTTDENGKYQFTGL---SNGTYSVEFSTPAGYTTTANVGTD 576  
QY 155 KTPPEKKEVTTKANLIYADGKTQTAEPFGTPEATAEAYRYADLLAKENG-----KYTV 209  
DB 577 -----DAVSDGLTTTGVTKD-----ADNMTLDSGFYKTPKYSL 610  
QY 210 -----DVADKGYTLNIFAGKEKTPPEKKEVTTKANLIYADGKTQTAEPK 255  
DB 611 GDYVWYDNGKGRDSTENG-----IKGVVTLQNEKEV-----GTTET----- 651  
QY 256 GTFABATAEAYRYADLLAKENGKTYAD-LEBGGYTINIRFAGKVDKPE 303  
DB 652 -----DENGKRYFDNLDGSKY-----KVIFEP 674

RESULT 13  
US-10-645-655-2  
; Sequence 2, Application US/10645655  
; GENERAL INFORMATION:  
; APPLICANT: St. Gene III, Joseph W.  
; APPLICANT: Falkow, Stanley  
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fiehr, Hobbach, Test, Albritton & Herbert  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco

```

; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/10/645,655
; FILING DATE: 20-AUG-2003
; CLASSIFICATION DATA:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-645-655-2

```

```

Query Match      7.3%; Score 114.5; DB 6; Length 1394;
Best Local Similarity 24.7%; Pred. No. 1.1; Indels 49; Gaps 17;
Matches 74; Conservative 45; Mismatches 131;

QY 7 ETPETDSEBEVITKANLIIFANGSTOTAEFGTPEKATSEAYAY-ADTLKONGEYTV 65
Db 863 ETEETP-TSAEH-----RFTLVNGKLSG---QGTFQ-FTSSLFGYKSDKLSND---- 909
QY 66 DVADKGYTLNIFKAGKEKTPPEKKEVITKANLIYADGKTQTAEFGTPE--EATAEAYR 123
Db 910 --AEGDITLSVRNTGKE---PETLEQLTVLVS---KDNQPLSDKLFLENDHVDAGALR 961
QY 124 YADALKKONGEYTVADVADKGYTLNIFKAGKEKTPPEKKEVITKANLIYADGKTQTAEFK 183
Db 962 Y--KLVDNGEPLHNPIKEQELHNDLVRAEQA-----ERTLEAKQVEPTAKTQTGEPK 1013
QY 184 GTFEEATAEAY-----RYADLLAKENGYTVADVADKGYTLNIFKAGKEKTPPEP--K 233
Db 1014 VRSRAARAAPDPTLPDQSLNLALEAKQAEATAE-TQSKAKTKKVRSKRAVFSPLDQ 1072
QY 234 EVITKANLIYADGKTQTAEFGTPEKATSEAYAYR-YADLLAKENGYTVADLEDGYTIN 291
Db 1073 SLFALEAALEVIDAPQOSE--KDLRAQEAEARQKQKDLISRYNSALSSEL---SATVN 1126

```

```

RESULT 14
US-10-408-765A-1139
; Sequence 1139, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04

```

```

; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1139
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1139

Query Match      7.2%; Score 113; DB 6; Length 1388;
Best Local Similarity 21.6%; Pred. No. 1.4; Indels 102; Gaps 14;
Matches 76; Conservative 40; Mismatches 134;

QY 5 KEETPE-----TPETDSE-----EVITKANLI-----FANGSTOTA- 36
Db 1022 REKTPVIDATEEIDLEETEREVSQENGLSEVPLGEMETDLKATGRDSPRGKTPEVI 1081
QY 37 ----EFKGTPEKATSEAYAYADTLK--KONGEYTVADVADKGYTLNIFKAGKEKTPPEPK 90
Db 1082 DAIBEIEIDLEETEREISPOENGLSEVPLGEMQTD-----LKATGREISPREKTP 1132
QY 91 EVITKANLIYADGKTQTAEFGTPEEATAEAYRYADALKKONGEYTVADVADKGYTLNIFK 150
Db 1133 EVI-----DATEEIDKLEETGRR-----EISPEENGPEEVPVDEMET-DLKT 1175
QY 151 AGKEKTPPEKKEVITKANLIYAD-----GKTQTAEFGTPEE- 188
Db 1176 TREGSSREKTRVIDAAEVIEDLEETEREISPOENGPEEVPKPKMET-DLKEIREEI 1234
QY 189 -----ATAEAYRYADLLAKENGYTVADVADK-----GYTLNIFKAGKEKTPPEP 232
Db 1235 SQREKVLAEFSAIREKEIDLKGTGRDIPIMEKVSQKMAVVEEMADLKTGKFNFRERG 1294
QY 233 KEEVTI---KANLIYADGKTQTAEFGTFAEATAEAYRYADLLAKENGYTA 281
Db 1295 SEECVTEEKVAELKQTKGTDISPNELEETSTSRQTDTHLMQSGSNDFSA 1346

RESULT 15
US-10-381-596A-2
; Sequence 2, Application US/10381596A
; GENERAL INFORMATION:
; APPLICANT: Biostaplo AB
; TITLE OF INVENTION: von Willebrand factor-binding proteins from
; FILE REFERENCE: 110059600
; CURRENT APPLICATION NUMBER: US/10/381,596A
; PRIOR FILING DATE: 2003-07-02
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2060
; TYPE: PRT
; ORGANISM: Staphylococcus lugdunensis
US-10-381-596A-2

Query Match      7.2%; Score 113; DB 6; Length 2060;
Best Local Similarity 20.2%; Pred. No. 2.2; Indels 36; Gaps 12;
Matches 62; Conservative 54; Mismatches 155;

QY 8 TPETPETDSEBEVITKANLIIFANGSTOTAEFGTPEKATSEAYAYADTLKONGEYTV-VD 66
Db 1299 TYEQIPNDAPQETPVALEV-----TRYVDSEGVQETEGTHDAPGIIADKWYTGQT 1352
QY 67 VADKGYTLNIFKAGKEKTPPEKKEVITKANLI-YADGKTQTAEFGTPEEATAEAYRYA 125
Db 1353 AAENGITTHVYQRIQSEIPNEAPQETPVALEVTQVDS-----EGNEVQETEGTHDA 1405
QY 126 DALKKONGEYTV-VDVADKGYTLNIFKAGKEKTPPEKKEVITKANLIYADGKTQTAEFGK 184
Db 1406 PGIIGDKWQYTGQTTEDGITTHYQRIQSEIPNEAPQETPVALEV-----TRYVDSEG 1459

```



Qy	185	TPEBTAAYRYADLLAKENGKKT--VDVADKGYYTLNIRPAGKXBTPBPBKBEVTIKANL	242
Dd	1460	NEVOETEETHOPPSIIGDKWQYGTQTTTAD-GITTVVYRIQSSEIPNEAPKETPIQLEV	1518
Qy	243	I-YADGKTQTBPKGTFPAETAAYRYADLLAKENGYTAD-LEDGGYINII--RPAGKK	298
Dd	1519	TRYVDG-----EGNEVOETEEGTHHAPGIIGDKWQYTQT <sup>7</sup> TTESGITTHVVYRQSEI	1571
Qy	299	VDEKPPEE	305
Dd	1572	PNEAPOSE	1578

Search completed: September 3, 2003, 11:14:30  
Job time : 14.3816 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:17 ; Search time 48.3288 Seconds  
(without alignments)  
236.470 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_153\_224

Perfect score: 369

Sequence: 1 KKKTPPEPKBVTIKANLIY.....GKYTVVDVADKGYTLNIKIFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	72	21	AA198539
2	369	100.0	305	14	AA198539
3	369	100.0	434	14	AA198539
4	358.5	97.2	467	22	AA198539
5	352	95.4	72	21	AA198538
6	352	95.4	367	21	AA198538
7	339	91.9	72	21	AA198540
8	315	85.4	75	21	AA198544
9	315	85.4	291	14	AA198544

10	315	85.4	1027	14	AA198539
11	315	85.4	1027	14	AA198539
12	301.5	81.7	71	21	AA198542
13	299	81.0	76	21	AA198537
14	293	79.4	82	21	AA198536
15	292	79.1	74	21	AA198543
16	286	77.5	82	21	AA198545
17	286	77.5	82	21	AA198548
18	284	77.0	82	21	AA198546
19	284	77.0	182	18	AA198547
20	284	77.0	182	20	AA198547
21	280	75.9	82	21	AA198547
22	276	74.8	178	18	AA198548
23	276	74.8	178	20	AA198548
24	276	74.8	198	18	AA198548
25	276	74.8	198	20	AA198548
26	276	74.8	342	18	AA198548
27	276	74.8	342	20	AA198548
28	276	74.8	482	20	AA198548
29	276	74.8	495	18	AA198548
30	276	74.8	495	20	AA198548
31	270.5	73.3	71	21	AA198541
32	80.5	21.8	395	22	ABG11860
33	79.5	21.5	463	22	ABG23272
34	67.5	18.3	368	22	ABG1686
35	67	18.2	450	22	AA198543
36	67	18.2	490	20	AA198543
37	67	18.2	544	17	AA198543
38	67	18.2	544	19	AA198543
39	67	18.2	608	22	AA198543
40	67	18.2	695	19	AA198543
41	67	18.2	695	21	AA198543
42	67	18.2	695	21	AA198543
43	67	18.2	924	12	AA198543
44	67	18.2	924	14	AA198543
45	67	18.2	924	14	AA198543

ALIGNMENTS

RESULT 1  
AA198539  
ID AA198539 standard; Protein; 72 AA.  
AC AA198539;  
XX  
XX  
DT 20-JUL-2000 (first entry)  
DE Peptostreptococcus strain 312 protein L domain B3 protein sequence.  
KW Immunoglobulin light chain binding protein; PpL; protein L;  
KW Peptostreptococcus; human immunoglobulin kappa chain;  
KW immunofluorescence chromatography.  
XX  
OS Peptostreptococcus sp.  
XX  
XX  
PN WO200015803-A1.  
XX  
PD 23-MAR-2000.  
XX  
PF 14-SEP-1999; 99WO-GB03048.  
XX  
PR 14-SEP-1998; 98GB-0019998.  
XX  
PR 26-APR-1999; 99GB-0009578.  
XX  
XX (ACTI-) ACTINOVA LTD.  
XX  
XX Gore MG, Beckingham JA, Roberts SE;  
XX WPI; 2000-271441/23.  
XX N-PSDB; AAA08428.

PT New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
XX more at pH8 with respect to human immunoglobulin kappa-chain -  
PS Disclosure; Page 43; 56pp; English.  
XX The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (P1) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. P1 is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC peptide sequence of protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
SQ Sequence 72 AA;  
Query Match 100.0%; Score 369; DB 21; Length 72;  
Best Local Similarity 100.0%; Pred. No. 9.4e-37;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKVTVDVA 60  
DB 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKVTVDVA 60  
QY 61 DKGYTLINIKFAG 72  
DB 61 DKGYTLINIKFAG 72  
RESULT 2  
ID AAR42993  
AC AAR42993 standard; Protein; 305 AA.  
XX AAR42993;  
XX  
DT 25-MAR-2003 (updated)  
DT 16-MAY-1994 (first entry)  
DE Immunoglobulin light chain binding protein (Protein L).  
XX Immunoglobulin; light chain; binding; identification; purification;  
KW separation.  
XX E. coli LE392/pHDL, DSM 7054.  
OS  
XX  
FH Key Location/Qualifiers  
FT Domain 5..305  
FT /label= B1 immunoglobulin light chain binding  
FT domain.  
FT Domain 81..305  
FT /label= B2 immunoglobulin light chain binding  
FT domain.  
FT Domain 153..305  
FT /label= B3 immunoglobulin light chain binding  
FT domain.  
FT Domain 225..305  
FT /label= B4 immunoglobulin light chain binding  
FT domain.  
FT Domain 297..305  
FT /label= B5 immunoglobulin light chain binding  
FT domain.  
XX WO9322342-A1.  
XX  
XX 11-NOV-1993.  
XX  
XX 28-APR-1993; 93WO-SE00375.  
XX  
XX 28-APR-1992; 92SE-0001331.  
XX  
XX (HIGH-) HIGHTECH RECEPTOR AB.  
XX  
XX Bjoerck L, Sjoerbring U;

XX WPI; 1993-368722/46.  
DR N-PSDB; AAQ50452.  
XX  
PT New protein L binding light chains of all immunoglobulin classes  
PT - for binding purifying and identifying immunoglobulin, also  
PT related DNA, vectors and host cells  
XX  
PS Claim 1; Page 36; 71pp; English.  
XX The protein (Protein L) is capable of binding to immunoglobulin G  
CC light chains. It is useful for binding, separating (purifying) and  
CC identifying immunoglobulin and for removing immunoglobulin molecules  
CC from serum. Hybrid proteins of the L protein can bind all human  
CC immunoglobulin classes and many immunoglobulins from other species.  
CC They are highly soluble and retain their binding activity at high  
CC temperatures over a pH range of 3-10. They can be immobilised  
CC without loss of activity.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 305 AA;  
Query Match 100.0%; Score 369; DB 14; Length 305;  
Best Local Similarity 100.0%; Pred. No. 6e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKVTVDVA 60  
DB 153 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKVTVDVA 212  
QY 61 DKGYTLINIKFAG 72  
DB 213 DKGYTLINIKFAG 224  
RESULT 3  
ID AAR42994  
AC AAR42994 standard; Protein; 434 AA.  
XX AAR42994;  
XX  
DT 25-MAR-2003 (updated)  
DT 16-MAY-1994 (first entry)  
DE Sequence encoding immunoglobulin light chain binding protein.  
XX Immunoglobulin; light chain; binding; identification; purification;  
KW separation; ss.  
XX E. coli L392/pHDLG, DSM 7055.  
OS  
XX  
FH Key Location/Qualifiers  
FT Domain 5..305  
FT /label= B1 immunoglobulin light chain binding  
FT domain.  
FT Domain 81..305  
FT /label= B2 immunoglobulin light chain binding  
FT domain.  
FT Domain 153..305  
FT /label= B3 immunoglobulin light chain binding  
FT domain.  
FT Domain 225..305  
FT /label= B4 immunoglobulin light chain binding  
FT domain.  
FT Domain 297..305  
FT /label= B5 immunoglobulin light chain binding  
FT domain.  
FT Domain 309..434  
FT /label= C1 immunoglobulin heavy chain binding  
FT domain.  
FT Domain 364..434  
FT /label= D intermediate immunoglobulin heavy  
FT chain binding domain.

```

PT Domain 379..434
FT /label= C2 immunoglobulin heavy chain binding
XX domain.
XX WO322342-A1.
XX 11-NOV-1993.
XX 28-APR-1993; 93WO-SE00375.
XX 28-APR-1992; 92SE-0001331.
XX (HIGH-) HIGHTECH RECEPTOR AB.
XX Bjoerck L, Sjoerbring U;
XX WPI; 1993-368722/46.
XX P-PSDB; AAR42994.
XX New protein L binding light chains of all immunoglobulin classes
PT - for binding purifying and identifying immunoglobulin, also
PT related DNA, vectors and host cells
XX Claim 6; Page 39-40; 71pp; English.
XX Protein L (AAR42993) is capable of binding to immunoglobulin G light
CC chains. It is useful for binding, separating (purifying) and
CC identifying immunoglobulin and for removing immunoglobulin molecules
CC from serum. This is the coding sequence of one hybrid protein of
CC the L protein. The hybrid proteins can bind all human
CC immunoglobulin classes and many immunoglobulins from other species.
CC They are highly soluble and retain their binding activity at high
CC temperatures over a pH range of 3-10. They can be immobilised
CC without loss of activity.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 434 AA;

Query Match 100.0%; Score 369; DB 14; Length 434;
Best Local Similarity 100.0%; Pred. No. 9.4e-36;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTFTFEATAEAYRYADLLAKENGKKTVDVA 60
Db 153 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTFTFEATAEAYRYADLLAKENGKKTVDVA 212

QY 61 DKGYTLNIKFPAG 72
Db 213 DKGYTLNIKFPAG 224

RESULT 4
ID AAB31372 standard; Protein; 467 AA.
XX AC
XX AAB31372;
XX 20-APR-2001 (first entry)
XX DE Amino acid sequence of protein L/CBD cex/ER retaining peptide fusion.
XX KW Protein production; food processing; protein antibiotic; feed enzyme;
XX KW protein L: CBD cex protein; cell signal peptide.
XX OS Synthetic.
XX WO200077174-A1.
XX 21-DEC-2000.
XX 07-JUN-2000; 2000WO-IL00330.
XX 10-JUN-1999; 99US-0329234.

```

---

```

XX (CBT-) CBD TECHNOLOGIES LTD.
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Shani Z, Shoeyov O;
XX WPI; 2001-112219/12.
DR N-PSDB; AAF24730.
XX Expressing and isolating recombinant protein in a plant, useful for
PT producing large quantities of recombinant proteins, by expressing a
PT fusion protein including a cellulose binding peptide fused to a
PT recombinant protein
XX Example; Fig 2a; 87pp; English.
XX The specification describes a method for expressing and isolating a
CC recombinant protein in a plant. The method comprising expressing a
CC fusion protein including the recombinant protein and a cellulose
CC binding peptide fused to it, where the fusion protein is
CC compartmentalised and sequestered within plant cells, plant derived
CC tissue or cultured plant cells. The method is useful for obtaining large
CC quantities of the recombinant proteins and protein products in a simple
CC and cost-effective manner. Recombinant proteins may be used commercially,
CC such as in the food processing industry, e.g. glucoamylases and glucose
CC isomerases are used for converting starch to high fructose corn syrup,
CC pectinases for the hydrolysis of high molecular weight proteins and in
CC manufacturing leather or alcoholic beverages, pectinesterases for
CC pectin hydrolysis in food industry, lipases for cleaving ester linkage
CC in triglycerides, and for effluent treatment. The recombinant proteins
CC may further be used to produce protein antibiotics, which can be used
CC in healing processes, and to produce animal feed enzymes. The present
CC sequence represents a fusion protein of the invention, and comprises a
CC fusion of a cell signal peptide, protein L, CBD cex, and an endoplasmic
CC reticulum retaining peptide.
XX SQ Sequence 467 AA;

Query Match 97.2%; Score 358.5; DB 22; Length 467;
Best Local Similarity 98.6%; Pred. No. 1.9e-34;
Matches 72; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTFTFEATAEAYRYADLLAKENGKKTVDV 59
Db 192 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTFTFEATAEAYRYADLLAKENGKKTVDV 251

QY 60 ADGYTLNIKFPAG 72
Db 252 ADGYTLNIKFPAG 264

RESULT 5
AAY82538
ID AAY82538 standard; Protein; 72 AA.
XX AC
XX AAY82538;
XX 20-JUL-2000 (first entry)
XX DE Peptostreptococcus strain 312 protein L domain B2 protein sequence.
XX KW Immunoglobulin light chain binding protein; Ppb; protein L;
XX KW Peptostreptococcus; human immunoglobulin kappa chain;
XX KW immunoaffinity chromatography.
XX OS Peptostreptococcus sp.
XX WO200015803-A1.
XX 23-MAR-2000.
XX 14-SEP-1999; 99WO-GB03048.
XX

```

```

PR 14-SEP-1998; 98GB-0019998.
PR 26-APR-1999; 99GB-0009578.
XX (ACTI-) ACTINOVA LTD.
PI Gore MG, Beckingham JA, Roberts SE;
XX WPI; 2000-271441/23.
DR N-PSDB; AAA08427.
XX New modified immunoglobulin light chain binding protein, useful in
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or
PT more at pH8 with respect to human immunoglobulin kappa-chain -
XX Disclosure; Page 42; 56pp; English.
XX The present invention describes an immunoglobulin (Ig) light chain
CC binding protein (p1) which has been modified by one or more amino acid
CC substitutions such that the dissociation constant (Kd) of the protein
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. p1 is
CC useful in immunoaffinity chromatography. The present sequence is a
CC Peptostreptococcus protein L Ig light chain binding domain, which is
CC given in the disclosure of the present invention.
XX
SQ Sequence 72 AA;
Query Match 95.4%; Score 352; DB 21; Length 72;
Best Local Similarity 94.4%; Pred. NO. 1e-34;
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKVTVDVA 60
DB 1 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADALKKONGEYTVDDA 60
QY 61 DKGTYTLNKFAG 72
DB 61 DKGTYTLNKFAG 72
RESULT 6
AAB10432
ID AAB10432 standard; Protein; 367 AA.
XX
AC AAB10432;
XX
DT 01-DEC-2000 (first entry)
XX
DE Expression vector pSEX1114 protein G.
XX
KW Expression vector; antibody binding protein; monoclonal antibody; Neo-R;
KW B lymphocyte; myeloma cell; hybridoma; protein G; bla protein.
XX
OS Synthetic.
XX
PN DE19900635-A1.
XX
PD 13-JUL-2000.
XX
PF 11-JAN-1999; 99DE-1000635.
XX
PR 11-JAN-1999; 99DE-1000635.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Breitling F, Poustka A, Moldenhauer G;
XX
DR WPI; 2000-499832/45.
DR N-PSDB; AAA71428.
XX
XX Selecting monoclonal antibodies, by expressing them on the surface of
PT hybridomas attached to antibody-binding protein, then reaction with
PT antibody library -
XX
PS Disclosure; Page 44-45; 56pp; English.

```

---

```

PS Claim 16; Fig 1; 22pp; German.
XX
CC This invention describes a novel method for the selection of monoclonal
CC antibodies (MAB) which comprises (i) fusing B lymphocytes with myeloma
CC cells to produce antibody-producing hybridomas such that the antibodies
CC are presented at the surface of the hybridomas by an antibody-binding
CC protein (I); and (ii) binding the antibody to antigens (Ag). The
CC invention also describes antibody-binding proteins (I) that comprise a
CC combination of the signal peptide of a murine immunoglobulin (Ig) kappa
CC chain or a murine MHC (major histocompatibility complex) Class I k(k)
CC molecule; an antibody-binding site of proteins A, G, L or LG, and the
CC transmembrane domain of PDGFR (platelet-derived growth factor receptor)
CC or CD52. The method is used to select MAB with specificity for particular
CC antigens. MAB can be selected without separate culture of hybridomas,
CC and selection can be made against many antigens in a library. Optionally
CC on the basis of strength of affinity for a particular antigen. Complex
CC mixtures of hybridomas can be used for selection, reducing the time and
CC cost involved in MAB selection. This sequence represents the protein G
CC contained in the expression vector pSEX1114 which contains the protein G,
CC Neo-R and the bla protein described in the method of the invention.
XX
SQ Sequence 367 AA;
Query Match 95.4%; Score 352; DB 21; Length 367;
Best Local Similarity 94.4%; Pred. NO. 8.3e-34;
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKVTVDVA 60
DB 25 KEKTEPEPKKEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADALKKONGEYTVDDA 84
QY 61 DKGTYTLNKFAG 72
DB 85 DKGTYTLNKFAG 96
RESULT 7
AAY82540
ID AAY82540 standard; Protein; 72 AA.
XX
AC AAY82540;
XX
DT 20-JUL-2000 (first entry)
XX
DE Peptostreptococcus strain 312 protein L domain B4 protein sequence.
XX
KW Immunoglobulin light chain binding protein; PpL; protein L;
KW Peptostreptococcus; human immunoglobulin kappa chain;
KW immunoaffinity chromatography.
XX
OS Peptostreptococcus sp.
XX
PN WO200015803-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-GB03048.
XX
PR 14-SEP-1998; 98GB-0019998.
PR 26-APR-1999; 99GB-0009578.
XX
PA (ACTI-) ACTINOVA LTD.
XX
PI Gore MG, Beckingham JA, Roberts SE;
XX
DR WPI; 2000-271441/23.
DR N-PSDB; AAA08429.
XX
XX New modified immunoglobulin light chain binding protein, useful in
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or
PT more at pH8 with respect to human immunoglobulin kappa-chain -
XX
PS Disclosure; Page 44-45; 56pp; English.

```

XX The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (p1) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. p1 is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.

XX  
SQ Sequence 72 AA;

Query Match 91.9%; Score 339; DB 21; Length 72;  
Best Local Similarity 90.3%; Pred. No. 3.7e-33; Mismatches 4; Indels 0; Gaps 0;  
Matches 65; Conservative 3;

QY 1 KEKTEPEKPEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKENGKKTVDVA 60  
DB 1 KEKTEPEKPEVTIKANLIYADGKTQTAEFKGTFAETAEAYRYADLLAKENGKKTADLE 60

QY 61 DGYTLNIRFAG 72  
DB 61 DGYTLNIRFAG 72

RESULT 8  
AAY82544  
ID AAY82544 standard; Protein; 75 AA.  
XX AC AAY82544;  
XX  
XX 20-JUL-2000 (first entry)  
XX  
XX DE Peptostreptococcus strain 3316 protein L domain C4 protein sequence.  
XX  
XX KW Immunoglobulin light chain binding protein; PpL; protein L;  
KW Peptostreptococcus; human immunoglobulin kappa chain;  
KW immunoaffinity chromatography.  
XX OS Peptostreptococcus sp.  
XX  
XX PN WO200015803-A1.  
XX  
XX PD 23-MAR-2000.  
XX  
XX PF 14-SEP-1999; 99WO-GB03048.  
XX  
XX PR 14-SEP-1998; 98GB-0019998.  
XX  
XX PR 26-APR-1999; 99GB-0009578.  
XX  
XX PA (ACTI-) ACTINOVA LTD.  
XX  
XX PI Gore MG, Beckingham JA, Roberts SE;  
XX  
XX WPI; 2000-271441/23.  
XX  
XX DR N-PSDB; AAA08433.  
XX  
XX New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
PT more at pH8 with respect to human immunoglobulin kappa-chain -  
XX  
XX PS Disclosure; Page 49-50; 56pp; English.  
XX  
XX CC The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (p1) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. p1 is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.

XX  
SQ Sequence 75 AA;

Query Match 85.4%; Score 315; DB 21; Length 75;

Best Local Similarity 85.7%; Pred. No. 3e-30; Mismatches 5; Indels 0; Gaps 0;  
Matches 60; Conservative 5;

QY 3 KTEPEKPEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKENGKKTVDVADK 62  
DB 5 ETPPEKPEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKVNGEYTTADLEDG 64

QY 63 GYTINIRFAG 72  
DB 65 GYTINIRFAG 74

RESULT 9  
AAR42204  
ID AAR42204 standard; Protein; 291 AA.  
XX AC AAR42204;  
XX  
XX 25-MAR-2003 (updated)  
DT 18-MAY-1994 (first entry)  
XX  
XX Immunoglobulin binding protein derived from protein L.  
XX  
XX KW Peptide; immunoglobulin; binding; analysis; purification; ELISA;  
KW enzyme linked immunosorbent assay.  
XX OS Synthetic.  
XX  
XX PN WO9322439-A1.  
XX  
XX PD 11-NOV-1993.  
XX  
XX PF 07-MAY-1993; 93WO-GB00950.  
XX  
XX PR 07-MAY-1992; 92GB-0009804.  
XX  
XX PR 24-DEC-1992; 92GB-0026928.  
XX  
XX PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.  
XX  
XX PI Atkinson A, Duggleby CJ, Murphy JP, Trowern AR;  
XX  
XX WPI; 1993-368798/46.  
XX  
XX DR N-PSDB; AAQ50947.  
XX  
XX New immunoglobulin binding proteins derived from Protein L -  
PT which bind immunoglobulin kappa light chains but not albumin or  
PT cell walls  
XX  
XX PS Claim 12; Figure 2; 28pp; English.  
XX  
XX CC The synthetic immunoglobulin binding proteins derived from protein  
CC L comprise repeated sequences from protein L which bind  
CC immunoglobulin kappa light chains. They can be used in protein  
CC analysis, purification procedures and other biochemical processes e.  
CC g. ELISA. The synthetic molecules are of particular advantage if  
CC they are free of regions in protein L which exhibit albumin and cell  
CC wall binding.  
XX  
XX CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX SQ Sequence 291 AA;

Query Match 85.4%; Score 315; DB 14; Length 291;  
Best Local Similarity 85.7%; Pred. No. 1.7e-29; Mismatches 5; Indels 0; Gaps 0;  
Matches 60; Conservative 5;

QY 3 KTEPEKPEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKENGKKTVDVADK 62  
DB 221 ETPPEKPEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKVNGEYTTADLEDG 280

QY 63 GYTINIRFAG 72  
DB 281 GYTINIRFAG 290

```
RESULT 10
AAR42203
ID AAR42203 standard; Protein; 1027 AA.
XX
AC AAR42203;
XX
DT 25-MAR-2003 (updated)
DT 18-MAY-1994 (first entry)
XX
DE Protein L.
XX
KW Peptide; immunoglobulin; binding; analysis; purification; ELISA;
KW enzyme linked immunoabsorbant assay.
XX
OS Peptococcus magnus.
XX
FH Key Location/Qualifiers
FT Peptide 36..59
FT /label= Signal sequence.
FT Protein 60..968
FT /label= Mature protein L.
XX
PN WO9322439-A1.
XX
PD 11-NOV-1993.
XX
PF 07-MAY-1993; 93WO-GB00950.
XX
PR 07-MAY-1992; 92GB-0009804.
PR 24-DEC-1992; 92GB-0026928.
XX
PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX
PI Atkinson A, Duggleby CJ, Murphy JP, Trowern AR;
DR WPI; 1993-368798/46.
DR N-PSDB; AAQ0946.
XX
PT New immunoglobulin binding proteins derived from Protein L -
PT which bind immunoglobulin kappa light chains but not albumin or
PT cell walls
XX
PS Disclosure; Figure 1; 28pp; English.
XX
CC The synthetic immunoglobulin binding proteins derived from protein
CC L correspond to the repeated sequences in protein L which bind
CC immunoglobulin kappa light chains. They can be used in protein
CC analysis, purification procedures and other biochemical processes e.
CC g. ELISA. The synthetic molecules are of particular advantage if
CC they are free of regions in protein L which exhibit albumin and cell
CC wall binding.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1027 AA;
Query Match 85.4%; Score 315; DB 14; Length 1027;
Best Local Similarity 85.7%; Pred. No. 8.6e-29;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 3 KTEPEPKKEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKENGKYYTDVADK 62
DB 503 ETPEPKKEVTIKVNLIFADGKTQTAEFKGTFEETAEAYRYADLLAKVNGEYTDLEDG 562
QY 63 GYTLNIKFPAG 72
DB 563 GTINIKFPAG 572
RESULT 11
AAR43699
ID AAR43699 standard; Protein; 1027 AA.
XX
```

```
AC AAR43699;
XX
DT 25-MAR-2003 (updated)
DT 18-MAY-1994 (first entry)
XX
DE Protein L.
XX
KW Protein; immunoglobulin; binding; immobilisation; light chains;
KW antibodies; diagnosis; pharmaceutical; ss.
XX
OS Peptococcus magnus.
XX
FH Key Location/Qualifiers
FT Peptide 36..59
FT /label= Signal sequence.
FT Protein 60..968
FT /label= Mature protein L.
XX
PN WO9322438-A1.
XX
PD 11-NOV-1993.
XX
PF 07-MAY-1993; 93WO-GB00949.
XX
PR 07-MAY-1992; 92GB-0009804.
XX
PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX
PI Atkinson A, Duggleby CJ, Murphy JP, Trowern AR;
DR WPI; 1993-368797/46.
DR P-PSDB; AAR43699.
XX
PT Immunoglobulin binding polypeptide, protein L - used for prodn.
PT of pharmaceuticals and for immobilising antibodies e.g. on
PT columns, in diagnostic tests and in assays
XX
PS Claim 4; Figure 1; 29pp; English.
XX
CC Protein L forms a complex with immunoglobulin kappa light chain.
CC Purified protein can be used as a reagent for immobilising
CC antibodies e.g. on columns, in diagnostic tests and in assays. It
CC may also be used in the production of pharmaceuticals.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1027 AA;
Query Match 85.4%; Score 315; DB 14; Length 1027;
Best Local Similarity 85.7%; Pred. No. 8.6e-29;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 3 KTEPEPKKEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKENGKYYTDVADK 62
DB 503 ETPEPKKEVTIKVNLIFADGKTQTAEFKGTFEETAEAYRYADLLAKVNGEYTDLEDG 562
QY 63 GYTLNIKFPAG 72
DB 563 GTINIKFPAG 572
RESULT 12
AAR82542
ID AAR82542 standard; Protein; 71 AA.
XX
AC AAR82542;
XX
DT 20-JUL-2000 (first entry)
XX
DE Peptostreptococcus strain 3316 protein L domain C2 protein sequence.
XX
KW Immunoglobulin light chain binding protein; PpL; protein L;
KW Peptostreptococcus; human immunoglobulin kappa chain;
KW immunoaffinity chromatography.
```





XX The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (P1) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. P1 is  
CC useful in immunoaffinity chromatography. The present sequence  
CC represents an Ig light chain binding PpL construct derived from  
CC Peptostreptococcus sp. protein L.  
XX  
SQ Sequence 82 AA;

Query Match 79.4%; Score 293; DB 21; Length 82;  
Best Local Similarity 81.4%; Pred. No. 1.5e-27;  
Matches 57; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 KTPPEPKKEVTIKANLIYADGKTOTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVADK 62  
DB 12 ETPPEPKKEVTIKVNLIFADGKIQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVADK 71  
QY 63 GYTLNLIKFPAG 72  
DB 72 GNHMIKFPAG 81

RESULT 15  
AA82543  
ID AAY82543 standard; Protein; 74 AA.  
XX  
AC AAY82543;  
XX  
DT 20-JUL-2000 (first entry)  
XX  
DE Peptostreptococcus strain 3316 protein L domain C3 protein sequence.  
XX  
KW Immunoglobulin light chain binding protein; PpL; protein L;  
KW Peptostreptococcus; human immunoglobulin kappa chain;  
KW immunoaffinity chromatography.  
XX  
OS Peptostreptococcus sp.  
XX  
PN WO200015803-A1.  
XX  
PD 23-MAR-2000.  
XX  
PF 14-SEP-1999; 99WO-GB03048.  
XX  
PR 14-SEP-1998; 98GB-0019998.  
PR 26-APR-1999; 99GB-0009578.  
XX  
PA (ACTI-) ACTINOVIA LTD.  
XX  
PI Gore MG, Beckingham JA, Roberts SE;  
XX  
DR WPI; 2000-271441/23.  
DR N-PSDB; AAA08543.  
XX  
PT New modified immunoglobulin light chain binding protein, useful in  
PT immunoaffinity chromatography, has a dissociation constant of 400 nM or  
PT more at pH8 with respect to human immunoglobulin kappa-chain -  
XX  
PS Disclosure; Page 48; 56pp; English.  
XX

XX The present invention describes an immunoglobulin (Ig) light chain  
CC binding protein (P1) which has been modified by one or more amino acid  
CC substitutions such that the dissociation constant (Kd) of the protein  
CC with respect to human Ig kappa-chain is 400 nM or more at pH8. P1 is  
CC useful in immunoaffinity chromatography. The present sequence is a  
CC Peptostreptococcus protein L Ig light chain binding domain, which is  
CC given in the disclosure of the present invention.  
XX  
SQ Sequence 74 AA;

Query Match 79.1%; Score 292; DB 21; Length 74;

Best Local Similarity 80.0%; Pred. No. 1.7e-27;  
Matches 56; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
QY 3 KTPPEPKKEVTIKANLIYADGKTOTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVADK 62  
DB 5 ETPPEPKKEVTIKVNLIFADGKIQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVADK 64  
QY 63 GYTLNLIKFPAG 72  
DB 65 GNHMIKFPAG 74

Search completed: September 3, 2003, 11:19:18  
Job time : 48.3288 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:16:02 ; Search time 15.7808 Seconds  
(without alignments)  
193.043 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_153\_224

Perfect score: 369  
Sequence: 1 KKTPEPKKEVTIKANLIY.....GKYTVADVADKGYTLNIPAG 72

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5A\_COMB.pdp.\*  
2: /cgm2\_6/ptodata/1/iaa/5B\_COMB.pdp.\*  
3: /cgm2\_6/ptodata/1/iaa/6A\_COMB.pdp.\*  
4: /cgm2\_6/ptodata/1/iaa/6B\_COMB.pdp.\*  
5: /cgm2\_6/ptodata/1/iaa/ECTUS\_COMB.pdp.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	305	2	US-08-795-475-1
2	369	100.0	434	2	US-08-795-475-3
3	315	85.4	75	3	US-08-446-137B-8
4	315	85.4	291	3	US-08-446-137B-4
5	315	85.4	1027	3	US-08-446-137B-2
6	301.5	81.7	71	3	US-08-446-137B-6
7	292	79.1	74	3	US-08-446-137B-7
8	284	77.0	182	3	US-08-828-741B-2
9	284	77.0	182	4	US-09-160-567-2
10	284	77.0	182	4	US-09-160-567-2
11	276	74.8	178	3	US-08-828-741B-13
12	276	74.8	178	4	US-09-160-567-13
13	276	74.8	178	4	US-09-160-567-13
14	276	74.8	198	3	US-08-828-741B-8
15	276	74.8	198	4	US-09-160-567-8
16	276	74.8	198	4	US-09-160-567-8
17	276	74.8	342	3	US-08-828-741B-6
18	276	74.8	342	4	US-09-160-567-6
19	276	74.8	342	4	US-09-160-567-6
20	276	74.8	495	3	US-08-828-741B-4
21	276	74.8	495	4	US-09-160-567-4
22	276	74.8	495	4	US-09-160-567-4
23	269.5	73.0	71	3	US-08-446-137B-5
24	67	18.2	490	4	US-09-252-149B-26
25	67	18.2	544	1	US-08-387-156-10
26	67	18.2	544	2	US-08-694-865-10
27	67	18.2	544	2	US-08-878-748-10

28	67	18.2	544	3	US-09-124-491-10	Sequence 10, Appl
29	67	18.2	544	4	US-09-383-912-10	Sequence 10, Appl
30	67	18.2	699	2	US-08-694-865-16	Sequence 16, Appl
31	67	18.2	699	3	US-09-124-491-16	Sequence 16, Appl
32	67	18.2	699	4	US-09-383-912-16	Sequence 16, Appl
33	67	18.2	924	3	US-08-619-812-8	Sequence 8, Appl
34	67	18.2	926	1	US-07-908-253-2	Sequence 2, Appl
35	67	18.2	926	1	US-08-455-970A-2	Sequence 2, Appl
36	67	18.2	926	1	US-08-387-156-6	Sequence 6, Appl
37	67	18.2	926	2	US-08-694-865-6	Sequence 6, Appl
38	67	18.2	926	2	US-08-878-748-6	Sequence 6, Appl
39	67	18.2	926	2	US-08-535-837-2	Sequence 2, Appl
40	67	18.2	926	3	US-09-124-491-6	Sequence 6, Appl
41	67	18.2	926	4	US-09-383-912-6	Sequence 6, Appl
42	67	18.2	926	6	5476657-3	Patent No. 5476657
43	67	18.2	936	1	US-08-455-970A-12	Sequence 12, Appl
44	67	18.2	943	1	US-08-455-970A-10	Sequence 10, Appl
45	67	18.2	951	1	US-08-455-970A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-08-795-475-1  
; Sequence 1, Application US/08795475  
; Patent No. 5965390  
; GENERAL INFORMATION:  
; APPLICANT: Bivrock, Lars  
; APPLICANT: Sjvbrck, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,475  
; FILING DATE: 11-FEB-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.402D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

Query Match 100.0%; Score 369; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 4.6e-39;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEPKKEVTIKANLIYADGKQTQAEFGKTFEATAEAYRYADLAKNGKTYDVA 60  
|||||

Db 153 KETPEEPKEVITKANLIYADGKTQTAEFKGTFFEATAEAYRYADLLAKENGKTYDVDA 212

QY 61 DKGYTLNLIKPFAG 72

Db 213 DKGYTLNLIKPFAG 224

## RESULT 2

US-08-795-475-3

; Sequence 3, Application US/08795475

; Patent No. 5965390

; GENERAL INFORMATION:

; APPLICANT: Bjvick, Lars

; APPLICANT: Sivbring, Ulf

; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/795,475

; FILING DATE: 11-FEB-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: McMasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 100084.402D1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 434 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055

US-08-795-475-3

Query Match 100.0%; Score 369; DB 2; Length 434;

Best Local Similarity 100.0%; Pred. No. 7.4e-39;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KETPEEPKEVITKANLIYADGKTQTAEFKGTFFEATAEAYRYADLLAKENGKTYDVDA 60

Db 153 KETPEEPKEVITKANLIYADGKTQTAEFKGTFFEATAEAYRYADLLAKENGKTYDVDA 212

QY 61 DKGYTLNLIKPFAG 72

Db 213 DKGYTLNLIKPFAG 224

## RESULT 3

US-08-446-137B-8

; Sequence 8, Application US/08446137B

; Patent No. 6162903

; GENERAL INFORMATION:

; APPLICANT: Trowern, Angus R.

; APPLICANT: Atkinson, Anthony

; APPLICANT: Murphy, Jonathan P.

; APPLICANT: Laurence, Oliver S.

; APPLICANT: Duggleby, Clive J.

; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED

; FROM L PROTEIN AND THEIR USES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,137B

; FILING DATE: 22-MAY-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: McMasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 100084.406

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 75 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-446-137B-8

Query Match 85.4%; Score 315; DB 3; Length 75;

Best Local Similarity 85.7%; Pred. No. 5.4e-33;

Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTEPEEKKEVITKANLIYADGKTQTAEFKGTFFEATAEAYRYADLLAKENGKTYDVADK 62

Db 5 ETPEEPKEVITKYNLIFADGKTQTAEFKGTFFEATAEAYRYADLLAKVNGEYTDADBDG 64

QY 63 GYTINIKFAG 72

Db 65 GYTINIKFAG 74

## RESULT 4

US-08-446-137B-4

; Sequence 4, Application US/08446137B

; Patent No. 6162903

; GENERAL INFORMATION:

; APPLICANT: Trowern, Angus R.

; APPLICANT: Atkinson, Anthony

; APPLICANT: Murphy, Jonathan P.

; APPLICANT: Laurence, Oliver S.

; APPLICANT: Duggleby, Clive J.

; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED

; FROM L PROTEIN AND THEIR USES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-137B-4

Query Match      85.4%; Score 315; DB 3; Length 291;
Best Local Similarity 85.7%; Pred. No. 3.3e-32;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTPPEPKKEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 62
Db 221 ETPPEPKKEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 280

QY 63 GYTLNPKFAG 72
Db 281 GYTLNPKFAG 290

RESULT 5
US-08-446-137B-2
; Sequence 2, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-446-137B-6

Query Match      81.7%; Score 301.5; DB 3; Length 71;
Best Local Similarity 83.3%; Pred. No. 2.6e-31;
Matches 60; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 KEKTPPEPKKEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 60
Db 1 KEK-PEEPKEEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 59

QY 61 DGVTNLIKFPAG 72
Db 60 DGVTNLIKFPAG 71

; MOLECULE TYPE: protein
US-08-446-137B-2

Query Match      85.4%; Score 315; DB 3; Length 1027;
Best Local Similarity 85.7%; Pred. No. 1.8e-31;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTPPEPKKEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 62
Db 503 ETPPEPKKEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 562

QY 63 GYTLNPKFAG 72
Db 563 GYTLNPKFAG 572

RESULT 6
US-08-446-137B-6
; Sequence 6, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Duggleby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 100084.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-446-137B-6

Query Match      81.7%; Score 301.5; DB 3; Length 71;
Best Local Similarity 83.3%; Pred. No. 2.6e-31;
Matches 60; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 KEKTPPEPKKEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 60
Db 1 KEK-PEEPKEEVTTKANLIYADGKTQTAEFGKTFEEATAYRYADLLAKENGKYYTDVADK 59

QY 61 DGVTNLIKFPAG 72
Db 60 DGVTNLIKFPAG 71
```

RESULT 7  
US-08-446-137B-7  
; Sequence 7, Application US/08446137B  
; Patent No. 6162903  
; GENERAL INFORMATION:  
; APPLICANT: Trowern, Angus R.  
; APPLICANT: Atkinson, Anthony  
; APPLICANT: Murphy, Jonathan P.  
; APPLICANT: Laurence, Oliver S.  
; APPLICANT: Duggieby, Clive J.  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; FROM L PROTEIN AND THEIR USES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,137B  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 100084.406  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-446-137B-7

Query Match 79.1%; Score 292; DB 3; Length 74;  
Best Local Similarity 80.0%; Pred. No. 4.5e-30;  
Matches 56; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
QY 3 KTEPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTVDVADK 62  
Db 5 ETPEPKKEVTIKVNLIFADGKIQTAEFGKTFEEATAKAYAYANLLAKENGGEYTDLEDG 64  
QY 63 GYTLINIKFAG 72  
Db 65 GNTINIKFAG 74

RESULT 8  
US-08-828-741B-2  
; Sequence 2, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCTION  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City

STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,741B  
FILING DATE: 26-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 182 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-828-741B-2  
Query Match 77.0%; Score 284; DB 3; Length 182;  
Best Local Similarity 76.1%; Pred. No. 1.6e-28;  
Matches 54; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 2 EKTPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTVDVAD 61  
Db 20 QAAPKDNTEEVTIKANLIFANGSTQTAEFGKTFEEATSEAYAYADTLKKDNGEYITVDVAD 79  
QY 62 KGYTLINIKFAG 72  
Db 80 KGYTLINIKFAG 90  
RESULT 9  
US-09-160-567-2  
; Sequence 2, Application US/09160567  
; Patent No. 6326179  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCTION  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```
RESULT 12
US-09-160-567-13
; Sequence 13, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-160-567-13

Query Match 74.8%; Score 276; DB 4; Length 178;
Best Local Similarity 84.1%; Pred. No. 1.6e-27;
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 EVVTKANLIYADGKTQTAEFGTPEEATAEAYRYADLLAKENGKYYTVDVADKGYTLNIK 69
DB 12 EVVTKANLIYADGKTQTAEFGTPEEATAEAYRYADLLAKENGKYYTVDVADKGYTLNIK 71
QY 70 FAG 72
DB 72 FAG 74

RESULT 13
US-09-710-299-13
; Sequence 13, Application US/09710299
; Patent No. 6521741
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-160-567-13

Query Match 74.8%; Score 276; DB 4; Length 178;
Best Local Similarity 84.1%; Pred. No. 1.6e-27;
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 EVVTKANLIYADGKTQTAEFGTPEEATAEAYRYADLLAKENGKYYTVDVADKGYTLNIK 69
DB 12 EVVTKANLIYADGKTQTAEFGTPEEATAEAYRYADLLAKENGKYYTVDVADKGYTLNIK 71
QY 70 FAG 72
DB 72 FAG 74

RESULT 13
US-09-710-299-13
; Sequence 13, Application US/09710299
; Patent No. 6521741
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/710,299
; FILING DATE: 09-NO. 6521741-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-710-299-13

Query Match 74.8%; Score 276; DB 4; Length 178;
Best Local Similarity 84.1%; Pred. No. 1.6e-27;
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 EVVTKANLIYADGKTQTAEFGTPEEATAEAYRYADLLAKENGKYYTVDVADKGYTLNIK 69
DB 12 EVVTKANLIYADGKTQTAEFGTPEEATAEAYRYADLLAKENGKYYTVDVADKGYTLNIK 71
QY 70 FAG 72
DB 72 FAG 74

RESULT 14
US-08-828-741B-8
; Sequence 8, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```



APPLICATION NUMBER: US/08/828,741B  
FILING DATE: 26-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-828-741B-8

Query Match 74.8%; Score 276; DB 3; Length 198;  
Best Local Similarity 84.1%; Pred. No. 1.9e-27;  
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 BEVTIKANLIYADGKTOTAEKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 69  
DB 32 BEVTIKANLIYADGKTOTAEKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 91  
QY 70 FAG 72  
DB 92 FAG 94

RESULT 15  
US-09-160-567-8  
Sequence 8, Application US/09160567  
Patent No. 6326179  
GENERAL INFORMATION:  
APPLICANT: Koentgen, Frank  
APPLICANT: Suesse, Gabriele M.  
APPLICANT: Tarlinton, David M.  
APPLICANT: Treutlein, Herbert R.  
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
TITLE OF INVENTION: PRODUCING SAME  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,567  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,741  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10591  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

LENGTH: 198 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-567-8

Query Match 74.8%; Score 276; DB 4; Length 198;  
Best Local Similarity 84.1%; Pred. No. 1.9e-27;  
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 BEVTIKANLIYADGKTOTAEKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 69  
DB 32 BEVTIKANLIYADGKTOTAEKGTPEEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 91  
QY 70 FAG 72  
DB 92 FAG 94

Search completed: September 3, 2003, 11:25:36  
Job time : 15.7808 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:23:22 ; Search time 26.6301 Seconds  
(without alignments)  
371.228 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_153\_224

Perfect score: 369

Sequence: 1 KEKTEPEKPEVTIKANLIY.....GKYTVADVADKGYTLNIKFAAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/ECT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/ECTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	72	10	US-09-808-212A-8
2	369	100.0	305	8	US-08-325-278-1
3	369	100.0	434	8	US-08-325-278-3
4	352	95.4	72	10	US-09-808-212A-6
5	339	91.9	72	10	US-09-808-212A-10
6	315	85.4	75	10	US-09-808-212A-18
7	301.5	81.7	71	10	US-09-808-212A-14
8	299	81.0	76	10	US-09-808-212A-4
9	293	79.4	82	10	US-09-808-212A-2
10	292	79.1	74	10	US-09-808-212A-16
11	284	77.0	182	12	US-10-345-618-2
12	276	74.8	178	12	US-10-345-618-13
13	276	74.8	198	12	US-10-345-618-8
14	276	74.8	342	12	US-10-345-618-6
15	276	74.8	482	12	US-10-345-618-16

16	276	74.8	495	12	US-10-345-618-4	Sequence 4, Appli
17	270.5	73.3	71	10	US-09-808-212A-12	Sequence 12, Appli
18	67	18.2	695	11	US-09-305-924-13	Sequence 13, Appli
19	67	18.2	953	11	US-09-884-696-3	Sequence 3, Appli
20	66	17.9	448	12	US-10-342-224-82	Sequence 82, Appli
21	66	17.9	1098	10	US-09-797-862-32	Sequence 32, Appli
22	65.5	17.8	271	15	US-10-138-505-30	Sequence 30, Appli
23	65.5	17.8	274	15	US-10-138-505-26	Sequence 26, Appli
24	65.5	17.8	836	10	US-09-858-525A-10	Sequence 10, Appli
25	65.5	17.8	871	10	US-09-858-525A-2	Sequence 2, Appli
26	63.5	17.2	279	14	US-10-059-964-60	Sequence 60, Appli
27	63.5	17.2	279	15	US-10-314-639-60	Sequence 60, Appli
28	63.5	17.2	414	9	US-09-039-927A-4	Sequence 4, Appli
29	63.5	17.2	1588	12	US-10-154-419-37	Sequence 37, Appli
30	63.5	17.2	1588	14	US-10-002-769-5	Sequence 5, Appli
31	63	17.1	284	11	US-09-056-019-3	Sequence 3, Appli
32	63	17.1	406	11	US-09-056-019-1	Sequence 1, Appli
33	63	17.1	428	11	US-09-056-019-24	Sequence 24, Appli
34	63	17.1	435	15	US-10-156-761-11066	Sequence 11066, A
35	63	17.1	446	12	US-10-254-395-9	Sequence 9, Appli
36	63	17.1	451	11	US-09-298-523B-67	Sequence 67, Appli
37	63	17.1	460	11	US-09-056-019-38	Sequence 38, Appli
38	63	17.1	511	11	US-09-298-523B-13	Sequence 13, Appli
39	63	17.1	513	11	US-09-298-523B-12	Sequence 12, Appli
40	63	17.1	655	11	US-09-056-019-2	Sequence 2, Appli
41	63	17.1	693	11	US-09-769-787-185	Sequence 185, App
42	61.5	16.7	136	15	US-10-138-505-8	Sequence 8, Appli
43	61.5	16.7	1179	15	US-10-156-761-13670	Sequence 13670, A
44	61	16.5	187	9	US-09-815-242-10591	Sequence 10591, A
45	61	16.5	841	9	US-09-815-242-5779	Sequence 5779, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-808-212A-8  
; Sequence 8, Application US/09808212A  
; Patent No. US20020137918A1  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-8

Query Match 100.0%; Score 369; DB 10; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1.6e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTEPEKPEVTIKANLIYADGKTQTAEFKGTFTTEATAEAYRYADLLAKENGYTVDA 60  
DB 1 KEKTEPEKPEVTIKANLIYADGKTQTAEFKGTFTTEATAEAYRYADLLAKENGYTVDA 60  
QY 61 DKGYTLNIKFAAG 72  
DB 61 DKGYTLNIKFAAG 72

##### RESULT 2

US-08-325-278-1  
; Sequence 1, Application US/08325278  
; Publication No. US20030027283A1  
; GENERAL INFORMATION:

APPLICANT: Bjvrck, Lars  
APPLICANT: Sjvbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McMasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCES/DOCKET NUMBER: 450023.401  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
US-08-325-278-1

Query Match 100.0%; Score 369; DB 8; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1e-34;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVA 60  
Db 153 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVA 212  
QY 61 DKGYTLNIKFPAG 72  
Db 213 DKGYTLNIKFPAG 224

RESULT 3  
US-08-325-278-3  
Sequence 3, Application US/08325278  
Publication No. US20030027283A1  
GENERAL INFORMATION:  
APPLICANT: Bjvrck, Lars  
APPLICANT: Sjvbring, Ulf  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278  
FILING DATE: 26-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McMasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCES/DOCKET NUMBER: 450023.401  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
US-08-325-278-3

Query Match 100.0%; Score 369; DB 8; Length 434;  
Best Local Similarity 100.0%; Pred. No. 1.6e-34;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVA 60  
Db 153 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVA 212  
QY 61 DKGYTLNIKFPAG 72  
Db 213 DKGYTLNIKFPAG 224

RESULT 4  
US-09-808-212A-6  
Sequence 6, Application US/09808212A  
Patent No. US20020137918A1  
GENERAL INFORMATION:  
APPLICANT: Gore, Michael Graham  
APPLICANT: Beckingham, Jennifer Ann  
APPLICANT: Roberts, Sian Eleri  
TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
FILE REFERENCE: 100084.414US  
CURRENT APPLICATION NUMBER: US/09/808,212A  
CURRENT FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-6

Query Match 95.4%; Score 352; DB 10; Length 72;  
Best Local Similarity 94.4%; Pred. No. 1.5e-33;  
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVA 60  
Db 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEEATAEAYRYADLLAKENGKKTVDVA 60  
QY 61 DKGYTLNIKFPAG 72  
Db 61 DKGYTLNIKFPAG 72

RESULT 5  
US-09-808-212A-10  
Sequence 10, Application US/09808212A  
Patent No. US20020137918A1

```

; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-10

Query Match      91.9%; Score 339; DB 10; Length 72;
Best Local Similarity 90.3%; Pred. No. 4.6e-32;
Matches 65; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTDVA 60
Db 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTDVA 60

Qy 61 DKGYYTLNPKFAG 72
Db 61 DGGYTNIRFAG 72

RESULT 6
US-09-808-212A-18
; Sequence 18, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-18

Query Match      85.4%; Score 315; DB 10; Length 75;
Best Local Similarity 85.7%; Pred. No. 2.9e-29;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 KTPPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTDVADK 62
Db 5 ETPEPKKEVTIKVNLIFADGKTQTAEFGKTFEEATAEAYRYADLLAKVNGEYTDLEDG 64

Qy 63 GYTLNPKFAG 72
Db 65 GYTLNPKFAG 74

RESULT 7
US-09-808-212A-14
; Sequence 14, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 2
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-14

Query Match      81.9%; Score 339; DB 10; Length 72;
Best Local Similarity 83.3%; Pred. No. 9.8e-28;
Matches 60; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTDVA 60
Db 1 KEKTPPEPKKEVTIKVNLIFADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYTDVA 59

Qy 61 DKGYYTLNPKFAG 72
Db 60 DGGNTINIKFAG 71

RESULT 8
US-09-808-212A-4
; Sequence 4, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-4

Query Match      81.0%; Score 299; DB 10; Length 76;
Best Local Similarity 77.6%; Pred. No. 2.1e-27;
Matches 59; Conservative 8; Mismatches 5; Indels 4; Gaps 1;

Qy 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTFEEATAEAYRYADLLAKENGKYYT 56
Db 1 KEETPTPTDSEEEVTIKANLIIFANGSTQTAEFGKTFEEATSEAYAYADTLKONGEYT 60

Qy 57 VDVAADKGYYTLNPKFAG 72
Db 61 VDVAADKGYYTLNPKFAG 76

RESULT 9
US-09-808-212A-2
; Sequence 2, Application US/09808212A
; Patent No. US20020137918A1
; GENERAL INFORMATION:
; APPLICANT: Gore, Michael Graham
; APPLICANT: Beckingham, Jennifer Ann
; APPLICANT: Roberts, Sian Eleri
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN
; FILE REFERENCE: 100084.414US
; CURRENT APPLICATION NUMBER: US/09/808,212A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Peptostreptococcus sp.
US-09-808-212A-2
```

```

RESULT 11
US-10-345-618-2
; Sequence 2, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:LHL protein
US-10-345-618-2

```

```

RESULT 13
US-10-345-618-8
; Sequence 8, Application US/10345618
; Publication No. US20030148484A1
; General INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBIOTIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 198
; TYPE: PRT

```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:LHL.seq
; OTHER INFORMATION: protein sequence
US-10-345-618-8

Query Match          74.8%; Score 276; DB 12; Length 198;
Best Local Similarity 84.1%; Pred. No. 3.2e-24;
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 EVVTIKANLIYADGKTQTAEFGKTPPEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 69
Db 32 EVVTIKANLIYADGKTQTAEFGKTPPEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 91
QY 70 FAG 72
Db 92 FAG 94

RESULT 14
US-10-345-618-6
; Sequence 6, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:TLHL protein
US-10-345-618-6

Query Match          74.8%; Score 276; DB 12; Length 342;
Best Local Similarity 84.1%; Pred. No. 6.4e-24;
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 EVVTIKANLIYADGKTQTAEFGKTPPEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 69
Db 176 EVVTIKANLIYADGKTQTAEFGKTPPEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 235
QY 70 FAG 72
Db 236 FAG 238

RESULT 15
US-10-345-618-16
; Sequence 16, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
```

```
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:ccMTLgL protein
US-10-345-618-16

Query Match          74.8%; Score 276; DB 12; Length 482;
Best Local Similarity 84.1%; Pred. No. 1e-23;
Matches 53; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 EVVTIKANLIYADGKTQTAEFGKTPPEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 69
Db 323 EVVTIKANLIYADGKTQTAEFGKTPPEATAEAYRYADLLAKENGYTVDVADKGYTLNIK 382
QY 70 FAG 72
Db 383 FAG 385

Search completed: September 3, 2003, 11:42:51
Job time : 27.6301 secs
```

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:19:27; Search time 213.041 Seconds  
(without alignments)  
294.149 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_153\_224

Perfect score: 369

Sequence: 1 KEKTPPEPKERVTKANLIY.....GKYTVADVADKGYTLNPKFAG 72

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgm2\_6/ptodata/1/paa/PCTUS COMB.pcp.\*
- 2: /cgm2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgm2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgm2\_6/ptodata/1/paa/US080 COMB.pcp.\*
- 5: /cgm2\_6/ptodata/1/paa/US081 COMB.pcp.\*
- 6: /cgm2\_6/ptodata/1/paa/US082 COMB.pcp.\*
- 7: /cgm2\_6/ptodata/1/paa/US083 COMB.pcp.\*
- 8: /cgm2\_6/ptodata/1/paa/US084 COMB.pcp.\*
- 9: /cgm2\_6/ptodata/1/paa/US085 COMB.pcp.\*
- 10: /cgm2\_6/ptodata/1/paa/US086 COMB.pcp.\*
- 11: /cgm2\_6/ptodata/1/paa/US087 COMB.pcp.\*
- 12: /cgm2\_6/ptodata/1/paa/US088 COMB.pcp.\*
- 13: /cgm2\_6/ptodata/1/paa/US089 COMB.pcp.\*
- 14: /cgm2\_6/ptodata/1/paa/US090 COMB.pcp.\*
- 15: /cgm2\_6/ptodata/1/paa/US091 COMB.pcp.\*
- 16: /cgm2\_6/ptodata/1/paa/US092 COMB.pcp.\*
- 17: /cgm2\_6/ptodata/1/paa/US093 COMB.pcp.\*
- 18: /cgm2\_6/ptodata/1/paa/US094 COMB.pcp.\*
- 19: /cgm2\_6/ptodata/1/paa/US095 COMB.pcp.\*
- 20: /cgm2\_6/ptodata/1/paa/US096 COMB.pcp.\*
- 21: /cgm2\_6/ptodata/1/paa/US097A COMB.pcp.\*
- 22: /cgm2\_6/ptodata/1/paa/US097B COMB.pcp.\*
- 23: /cgm2\_6/ptodata/1/paa/US098 COMB.pcp.\*
- 24: /cgm2\_6/ptodata/1/paa/US099A COMB.pcp.\*
- 25: /cgm2\_6/ptodata/1/paa/US099B COMB.pcp.\*
- 26: /cgm2\_6/ptodata/1/paa/US100 COMB.pcp.\*
- 27: /cgm2\_6/ptodata/1/paa/US101 COMB.pcp.\*
- 28: /cgm2\_6/ptodata/1/paa/US102 COMB.pcp.\*
- 29: /cgm2\_6/ptodata/1/paa/US103 COMB.pcp.\*
- 30: /cgm2\_6/ptodata/1/paa/US104 COMB.pcp.\*
- 31: /cgm2\_6/ptodata/1/paa/US60 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	72	23	US-09-808-212A-8
2	369	100.0	305	7	US-08-325-278-1

Sequence 8, Appli  
Sequence 1, Appli

3	369	100.0	305	7	US-08-325-278A-1	Sequence 1, Appli
4	369	100.0	305	7	US-08-325-278B-1	Sequence 1, Appli
5	369	100.0	434	7	US-08-325-278-3	Sequence 3, Appli
6	369	100.0	434	7	US-08-325-278A-3	Sequence 3, Appli
7	369	100.0	434	7	US-08-325-278B-3	Sequence 3, Appli
8	369	100.0	719	22	US-09-791-537-10210	Sequence 10210, A
9	369	100.0	719	22	US-09-791-537-96101	Sequence 96101, A
10	359.5	97.2	467	25	US-09-980-469-12	Sequence 12, Appli
11	352	95.4	72	23	US-09-808-212A-6	Sequence 6, Appli
12	352	95.4	367	23	US-09-889-182A-4	Sequence 4, Appli
13	339	91.9	72	23	US-09-808-212A-10	Sequence 10, Appli
14	315	85.4	75	23	US-09-808-212A-18	Sequence 18, Appli
15	315	85.4	291	8	US-08-446-137A-4	Sequence 4, Appli
16	315	85.4	992	22	US-09-791-537-88366	Sequence 88366, A
17	315	85.4	1027	7	US-08-331-637-2	Sequence 2, Appli
18	315	85.4	1027	8	US-08-446-137A-2	Sequence 2, Appli
19	315	85.4	1027	15	US-09-187-295-2	Sequence 2, Appli
20	301.5	81.7	71	23	US-09-808-212A-14	Sequence 14, Appli
21	299	81.0	76	23	US-09-808-212A-4	Sequence 4, Appli
22	299	81.0	78	22	US-09-791-537-22553	Sequence 22553, A
23	293	79.4	82	23	US-09-808-212A-2	Sequence 2, Appli
24	292	79.1	74	23	US-09-808-212A-16	Sequence 16, Appli
25	284	77.0	182	19	US-09-509-031-2	Sequence 2, Appli
26	284	77.0	182	23	US-09-820-048A-2	Sequence 2, Appli
27	276	74.8	178	19	US-09-509-031-13	Sequence 13, Appli
28	276	74.8	178	23	US-09-820-048A-13	Sequence 13, Appli
29	276	74.8	198	19	US-09-509-031-8	Sequence 8, Appli
30	276	74.8	198	23	US-09-820-048A-8	Sequence 8, Appli
31	276	74.8	342	19	US-09-509-031-6	Sequence 6, Appli
32	276	74.8	342	23	US-09-820-048A-6	Sequence 6, Appli
33	276	74.8	482	19	US-09-509-031-16	Sequence 16, Appli
34	276	74.8	495	19	US-09-509-031-4	Sequence 4, Appli
35	276	74.8	495	23	US-09-820-048A-4	Sequence 4, Appli
36	270.5	73.3	71	23	US-09-808-212A-12	Sequence 12, Appli
37	80.5	21.8	395	1	PCT-US01-08631-42219	Sequence 42219, A
38	80.5	21.8	395	19	US-09-538-092-857	Sequence 857, App
39	79.5	21.5	463	1	PCT-US01-08631-53631	Sequence 53631, A
40	72	19.5	891	18	US-09-451-320-2418	Sequence 2418, Ap
41	71.5	19.4	1378	31	US-60-167-217-8500	Sequence 8500, Ap
42	70.5	19.1	586	28	US-10-282-122A-53124	Sequence 53124, A
43	69	18.7	321	31	US-60-360-039-3708	Sequence 3708, Ap
44	69	18.7	1447	30	US-10-437-963-198658	Sequence 198658, A
45	68	18.4	414	21	US-09-708-427-20185	Sequence 20185, A

ALIGNMENTS

RESULT 1  
US-09-808-212A-8  
; Sequence 8, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 72  
; TYPE: PPT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-8

Query Match 100.0%; Score 369; DB 23; Length 72;  
Best Local Similarity 100.0%; Pred. No. 2.5e-36;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKERVTKANLIYADGKTQTAFKGTFTFEATPAAYRVLAKENGKTYDVA 60  
|||||

```
Db 1 KKTPEEPKEVTTKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 60
QY 61 DKGYTLNIKFPAG 72
Db 61 DKGYTLNIKFPAG 72

RESULT 2
US-08-325-278A-1
; Sequence 1, Application US/08325278
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McWaters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278-1

Query Match 100.0%; Score 369; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.9e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEVTTKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 60
Db 153 KKTPEEPKEVTTKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 212

QY 61 DKGYTLNIKFPAG 72
Db 213 DKGYTLNIKFPAG 224

RESULT 3
US-08-325-278A-1
; Sequence 1, Application US/08325278A
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278B
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
```

```
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278A
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 100084.402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 305 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-325-278A-1

Query Match 100.0%; Score 369; DB 7; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.9e-35;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKTPEEPKEVTTKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 60
Db 153 KKTPEEPKEVTTKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 212

QY 61 DKGYTLNIKFPAG 72
Db 213 DKGYTLNIKFPAG 224

RESULT 4
US-08-325-278B-1
; Sequence 1, Application US/08325278B
; GENERAL INFORMATION:
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278B
; FILING DATE: 26-Oct-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
```

REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 100084.402  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Escherichia coli LE392/pHDL, DSM 7054  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-08-325-278B-1

Query Match 100.0%; Score 369; DB 7; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1.9e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTPFEATAEAYRYADLLAKENGKTYVDVA 60  
DB 153 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTPFEATAEAYRYADLLAKENGKTYVDVA 212  
QY 61 DKGYTLNIKFPAG 72  
DB 213 DKGYTLNIKFPAG 224

RESULT 5  
US-08-325-278-3  
; Sequence 3, Application US/08325278  
; GENERAL INFORMATION:  
; APPLICANT: Bjvrck, Lars  
; APPLICANT: Sjovring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278  
; FILING DATE: 26-OCT-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcmasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 450023.401  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7055  
US-08-325-278-3

Query Match 100.0%; Score 369; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 3.1e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTPFEATAEAYRYADLLAKENGKTYVDVA 60  
DB 153 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTPFEATAEAYRYADLLAKENGKTYVDVA 212  
QY 61 DKGYTLNIKFPAG 72  
DB 213 DKGYTLNIKFPAG 224

RESULT 6  
US-08-325-278A-3  
; Sequence 3, Application US/08325278A  
; GENERAL INFORMATION:  
; APPLICANT: Bjvrck, Lars  
; APPLICANT: Sjovring, Ulf  
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/325,278A  
; FILING DATE: 26-Oct-1994  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E. R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 100084.402  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli LE392/pHDL, DSM 7055  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-08-325-278A-3

Query Match 100.0%; Score 369; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 3.1e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTPFEATAEAYRYADLLAKENGKTYVDVA 60  
DB 153 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGKTPFEATAEAYRYADLLAKENGKTYVDVA 212  
QY 61 DKGYTLNIKFPAG 72  
DB 213 DKGYTLNIKFPAG 224

RESULT 7  
US-08-325-278B-3

Sequence 3, Application US/08325278B  
GENERAL INFORMATION:  
APPLICANT: Bjorck, Lars  
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Fifth Avenue Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,278B  
FILING DATE: 26-Oct-1994  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 100084.402  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-08-325-278B-3  
Query Match 100.0%; Score 369; DB 7; Length 434;  
Best Local Similarity 100.0%; Pred. No. 3.1e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 60  
Db 153 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 212  
QY 61 DKGTYLNIKFPAG 72  
Db 213 DKGTYLNIKFPAG 224  
RESULT 8  
US-09-791-537-10210  
Sequence 10210, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent In version 3.0  
SEQ ID NO 10210  
LENGTH: 719  
TYPE: PRT  
ORGANISM: Peptostreptococcus magnus

US-09-791-537-10210  
Query Match 100.0%; Score 369; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 6.2e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 60  
Db 246 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 305  
QY 61 DKGTYLNIKFPAG 72  
Db 306 DKGTYLNIKFPAG 317  
RESULT 9  
US-09-791-537-96101  
Sequence 96101, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 96101  
LENGTH: 719  
TYPE: PRT  
ORGANISM: Peptostreptococcus magnus  
US-09-791-537-96101  
Query Match 100.0%; Score 369; DB 22; Length 719;  
Best Local Similarity 100.0%; Pred. No. 6.2e-35;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 60  
Db 246 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADLLAKENGKYYTDVA 305  
QY 61 DKGTYLNIKFPAG 72  
Db 306 DKGTYLNIKFPAG 317  
RESULT 10  
US-09-980-469-12  
Sequence 12, Application US/09980469  
GENERAL INFORMATION:  
APPLICANT: Shoseyov, Oded  
APPLICANT: Ziv, Shani  
TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND RECO  
FILE REFERENCE: 01/22924  
CURRENT APPLICATION NUMBER: US/09/980,469  
CURRENT FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 12  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: pUC19-cell-ProL1-cexNG-HDEL fusion encoded product  
US-09-980-469-12  
Query Match 97.2%; Score 358.5; DB 25; Length 467;  
Best Local Similarity 98.6%; Pred. No. 6.3e-34;  
Matches 72; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLL-AKENGKKTVDV 59  
DB 192 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAAKENGKKTVDV 251  
QY 60 ADKGYTLNKFAG 72  
DB 252 ADKGYTLNKFAG 264

RESULT 11  
US-09-808-212A-6  
; Sequence 6, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-6

Query Match 95.4%; Score 352; DB 23; Length 72;  
Best Local Similarity 94.4%; Pred. No. 2.8e-34;  
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKKTVDVA 60  
DB 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADALKKONGEYTTDVA 60

QY 61 DKGYTLNKFAG 72  
DB 61 DKGYTLNKFAG 72

RESULT 12  
US-09-889-182A-4  
; Sequence 4, Application US/09889182A  
; GENERAL INFORMATION:  
; APPLICANT: Breitling, Frank  
; TITLE OF INVENTION: SELECTION OF MONOCLONAL ANTIBODIES  
; FILE REFERENCE: 4121-126  
; CURRENT APPLICATION NUMBER: US/09/889,182A  
; CURRENT FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: pct/de00/00079  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 367  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-09-889-182A-4

Query Match 95.4%; Score 352; DB 23; Length 367;  
Best Local Similarity 94.4%; Pred. No. 2.7e-33;  
Matches 68; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKKTVDVA 60  
DB 25 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADALKKONGEYTTDVA 84

QY 61 DKGYTLNKFAG 72  
DB 85 DKGYTLNKFAG 96

## RESULT 13

US-09-808-212A-10  
; Sequence 10, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-10

Query Match 91.9%; Score 339; DB 23; Length 72;  
Best Local Similarity 90.3%; Pred. No. 1e-32;  
Matches 65; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKKTVDVA 60  
DB 1 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKKTADLE 60

QY 61 DKGYTLNKFAG 72  
DB 61 DGYTINIRFAG 72

## RESULT 14

US-09-808-212A-18  
; Sequence 18, Application US/09808212A  
; GENERAL INFORMATION:  
; APPLICANT: Gore, Michael Graham  
; APPLICANT: Beckingham, Jennifer Ann  
; APPLICANT: Roberts, Sian Eleri  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN  
; FILE REFERENCE: 100084.414US  
; CURRENT APPLICATION NUMBER: US/09/808,212A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Peptostreptococcus sp.  
US-09-808-212A-18

Query Match 85.4%; Score 315; DB 23; Length 75;  
Best Local Similarity 85.7%; Pred. No. 8.7e-30;  
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KKTPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKENGKKTVDVADK 62  
DB 5 ETPPEEPKEEVTIKANLIYADGKTQTAEFKGTPEATAEAYRYADLLAKNGEYTTADLE 64

QY 63 GYTLNKFAG 72  
DB 65 GYTINIKFAG 74

## RESULT 15

US-08-446-137A-4  
; Sequence 4, Application US/08446137A  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED  
; FROM PROTEIN L AND THEIR USES

```
;
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137A
; FILING DATE: 22-05-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MITCHARD, LEONARD C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 1418-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-137A-4

Query Match      85.4%; Score 315; DB 8; Length 291;
Best Local Similarity 85.7%; Pred. No. 5.8e-29;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      3 KTEPEPKKEVTIKANLIYADGKTQTAEFKGTFEETAEAYRYADLLAKENGYKTYDVADK 62
Db      221 ETPPEPKKEVTIKVNLIFADGKTQTAEFKGTFEETAEAYRYADLLAKVNGEYTTADLEDG 280

QY      63 GYTLNIKFPAG 72
Db      281 GYTINIKFPAG 290

Search completed: September 3, 2003, 11:40:08
Job time : 214.041 secs
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:20:12 ; Search time 10.1096 Seconds  
(without alignments)  
223.701 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_153\_224

Perfect score: 369

Sequence: 1 KEKTEPEKKEVTKANLIY.....GKYTVDVADKGYTLNIKIPAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 148256 seqs, 31410045 residues

Total number of hits satisfying chosen parameters: 148256

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA.New.\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.5	17.8	271	6	US-10-221-131-95
2	65.5	17.8	271	6	US-10-221-131-96
3	63.5	17.2	914	6	US-10-631-467-789
4	63.5	17.2	1019	1	PCT-US03-21510-62
5	63	17.1	693	6	US-10-333-120A-10
6	62.5	16.9	423	7	US-60-491-156-12
7	62.5	16.9	1849	6	US-10-637-544-2
8	61.5	16.7	136	6	US-10-221-131-91
9	61	16.5	930	5	US-09-200-6508-3
10	61	16.5	1166	5	US-09-200-6508-7
11	60.5	16.4	636	6	US-10-613-520-1638
12	60.5	16.4	637	5	US-09-897-516A-7157
13	60.5	16.4	638	6	US-10-613-520-1341
14	60.5	16.4	638	6	US-10-613-520-1356
15	60.5	16.4	713	6	US-10-467-534-97
16	60.5	16.4	2315	6	US-10-603-114-5434
17	60	16.3	291	6	US-10-640-833-4201
18	60	16.3	328	6	US-10-617-320-3318
19	60	16.3	728	6	US-10-467-534-81
20	59	16.0	242	5	US-09-674-546A-244
21	59	16.0	1646	6	US-10-640-833-4667
22	59	16.0	1653	6	US-10-617-320-4155
23	58	15.7	335	6	US-10-640-833-4355
24	58	15.7	342	6	US-10-617-320-3303
25	58	15.7	722	6	US-09-897-516A-6073
26	58	15.7	820	5	US-10-640-833-4219

27	58	15.7	834	6	US-10-617-320-4760
28	58	15.7	1684	6	US-10-408-765A-811
29	57.5	15.6	245	6	US-10-221-131-100
30	57.5	15.6	271	6	US-10-221-131-98
31	57.5	15.6	274	6	US-10-221-131-97
32	57.5	15.6	533	6	US-10-221-131-101
33	57.5	15.6	753	6	US-10-603-114-5022
34	57.5	15.6	1096	6	US-10-451-901-1
35	57.5	15.6	1262	6	US-10-293-244-1125
36	57.5	15.6	1289	6	US-10-293-244-3093
37	57	15.4	224	6	US-10-286-897-5675
38	57	15.4	224	6	US-10-258-898A-5675
39	57	15.4	311	6	US-10-286-897-2103
40	57	15.4	311	6	US-10-258-898A-2103
41	57	15.4	513	6	US-10-603-114-4981
42	57	15.4	977	6	US-10-293-244-1201
43	57	15.4	977	7	US-60-490-890-580
44	57	15.4	1003	6	US-10-293-244-3169
45	57	15.4	2073	7	US-60-490-890-980

ALIGNMENTS

RESULT 1

US-10-221-131-95  
; Sequence 95, Application US/10221131  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, NAOSHI  
; APPLICANT: TSUCHIYA, MASAYUKI  
; APPLICANT: OH-EDA, MASAYOSHI  
; APPLICANT: UNO, SHINSUKE  
; APPLICANT: KIKUCHI, YASUFUMI  
; TITLE OF INVENTION: POLYPEPTIDE INDUCING APOPTOSIS  
; FILE REFERENCE: 065678/0106  
; CURRENT APPLICATION NUMBER: US/10/221,131  
; CURRENT FILING DATE: 2003-02-11  
; PRIOR APPLICATION NUMBER: PCT/JP01/01912  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 09/523,095  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: JP 2000-115246  
; PRIOR FILING DATE: 2000-04-17  
; PRIOR APPLICATION NUMBER: JP 2000-321822  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 103  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 95  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Mus sp.  
; FEATURE:  
; OTHER INFORMATION: pscml. MABL1-scfv

Query Match 17.8%; Score 65.5; DB 6; Length 271;

Best Local Similarity 30.3%; Pred. No. 1.9;  
Matches 20; Conservative 9; Mismatches 24; Indels 13; Gaps 2;

QY	20	YADGKTQTAEPKG-----TFEATATAEVRYADLLAKEN-----GKYTVDVADKGYTL 66
Db	76	YNDGTYNEKFKGKATLTSEKSSAAVYMLSSLASDSAVTYCARGYYSDMDMGQGTTL 135
QY	67	NIKIPAG 72
Db	136	TVSSGG 141

RESULT 2

US-10-221-131-96  
; Sequence 96, Application US/10221131  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, NAOSHI

```
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; TITLE OF INVENTION: POLYPEPTIDE INDUCING APOPTOSIS
; FILE REFERENCE: 065678/0106
; CURRENT APPLICATION NUMBER: US/10/221,131
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 09/523,095
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: JP 2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP 2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 96
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: pCHOM1. MABL1-scFv
US-10-221-131-96

Query Match 17.8%; Score 65.5; DB 6; Length 271;
Best Local Similarity 30.3%; Pred. No. 1.9;
Matches 20; Conservative 9; Mismatches 24; Indels 24; Gaps 2;

QY 20 YADGKTQTAEFKG----TPEATAEAYRYADLLAKEN-----GKTYTDVADKGYTL 66
DB 73 YNDGTYKNEKFGKATLTSEKSSAAYMELSLASEDSAVYVCARGGYSDYDWDGQGTTL 132

QY 67 NIKFAG 72
DB 133 TVSSGG 138

RESULT 3
US-10-631-467-789
; Sequence 789, Application US/10631467
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 789
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-789

Query Match 17.2%; Score 63.5; DB 6; Length 914;
Best Local Similarity 40.5%; Pred. No. 15;
Matches 17; Conservative 5; Mismatches 19; Indels 19; Gaps 1;

QY 3 KTEPEKPEVTKANLIYADGKTQTAEFKGTPEATAEAYRY 44
DB 434 KTPSITSEAVPEAGLV-IDGKTLNAIFQGLEKFKLELTQY 474

RESULT 4
PCT-US03-21510-62
; Sequence 62, Application PC/TUS0321510
```

```
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MP21S AS MODIFIERS OF THE p21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-046C-PC
; CURRENT APPLICATION NUMBER: PCT/US03/21510
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US 60/394,795
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/401,739
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,010
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/437,158
; PRIOR FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 62
; LENGTH: 1019
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21510-62

Query Match 17.2%; Score 63.5; DB 1; Length 1019;
Best Local Similarity 40.5%; Pred. No. 17;
Matches 17; Conservative 5; Mismatches 19; Indels 1; Gaps 1;

QY 3 KTEPEKPEVTKANLIYADGKTQTAEFKGTPEATAEAYRY 44
DB 539 KTPSITSEAVPEAGLV-IDGKTLNAIFQGLEKFKLELTQY 579

RESULT 5
US-10-333-120A-10
; Sequence 10, Application US/10333120A
; GENERAL INFORMATION:
; APPLICANT: Janulczyk, Robert
; APPLICANT: Iannelli, Francesco
; APPLICANT: Sjöholm, Anders
; APPLICANT: Pozzi, Gianni
; APPLICANT: Björck, Lars
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 13519PCTUS
; CURRENT APPLICATION NUMBER: US/10/333,120A
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: PCT/EP01/08409
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: SE 0002728-4
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-333-120A-10

Query Match 17.1%; Score 63; DB 6; Length 693;
Best Local Similarity 40.4%; Pred. No. 12;
Matches 21; Conservative 5; Mismatches 20; Indels 6; Gaps 2;

QY 7 EPKEVTKANLIYADGK----TQTAEPKGTPEATAEAYRYADLLAKENGK 54
DB 238 EPRDQKIQAEVSEKQAEATRLKIKTDREAEAEAKRRAD--AKEQKG 287

RESULT 6
US-60-491-156-12
; Sequence 12, Application US/60491156
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliasof, Scott
```



;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
;; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 14303, 16816, 17827, 32620, 577,  
;; TITLE OF INVENTION: 619, 1423, 2158, 8263, 15402, 16209, 16386, 21165, 30911 OR  
;; TITLE OF INVENTION: 41897  
;; FILE REFERENCE: MPI03-098PIM  
;; CURRENT APPLICATION NUMBER: US/60/491,156  
;; CURRENT FILING DATE: 2003-07-30  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12  
;; LENGTH: 423  
;; TYPE: PRT  
;; ORGANISM: Homo Sapiens  
US-60-491-156-12

Query Match 16.9%; Score 62.5; DB 7; Length 423;  
Best Local Similarity 31.1%; Pred. No. 7.4;  
Matches 19; Conservative 12; Mismatches 23; Indels 7; Gaps 3;  
QY 1 KEKTPPEPKSEVTIKANLIYADGKTQTAEFKGTFFETATB---AYRYADLLAKENGKTYV 57  
Db 294 KAQLPKRELEIVVILGMEATG--MTCQARSSY--ITSEILLGWYRFTPLVLTLEDGFYEV 349  
QY 58 D 58  
Db 350 D 350

RESULT 7  
US-10-637-544-2  
;; Sequence 2, Application US/10637544  
;; GENERAL INFORMATION:  
;; APPLICANT: Chr. Hansen A/S  
;; TITLE OF INVENTION: Peptides with anti-hypertensive properties  
;; FILE REFERENCE: P1032US01  
;; CURRENT APPLICATION NUMBER: US/10/637,544  
;; CURRENT FILING DATE: 2003-08-11  
;; NUMBER OF SEQ ID NOS: 20  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 2  
;; LENGTH: 1849  
;; TYPE: PRT  
;; ORGANISM: Lactobacillus helveticus  
US-10-637-544-2

Query Match 16.9%; Score 62.5; DB 6; Length 1849;  
Best Local Similarity 33.8%; Pred. No. 47;  
Matches 25; Conservative 8; Mismatches 30; Indels 11; Gaps 4;  
QY 1 KEKTPPEPKSEVTIKANL---IYADGKTQ---TAEFKGTFFETATBAYRYADLLAKENG 53  
Db 1590 KENQPSDEKELAKAKENLQAKVDAGEKDLDKYTADSKKDFNDALKA---KDVLDKVA 1646  
QY 54 KYTVVDVADKGYTLN 67  
Db 1647 K-LADLQDAAKALD 1659

RESULT 8  
US-10-221-131-91  
;; Sequence 91, Application US/10221131  
;; GENERAL INFORMATION:  
;; APPLICANT: FUKUSHIMA, NAOSHI  
;; APPLICANT: TSUCHIYA, MASAYUKI  
;; APPLICANT: OH-EDA, MASAYOSHI  
;; APPLICANT: UNO, SHINSUKE  
;; APPLICANT: KIKUCHI, YASUFUMI  
;; TITLE OF INVENTION: POLYPEPTIDE INDUCING APOPTOSIS  
;; FILE REFERENCE: 065678/0106  
;; CURRENT APPLICATION NUMBER: US/10/221,131  
;; CURRENT FILING DATE: 2003-02-11  
;; PRIOR APPLICATION NUMBER: PCT/JP01/01912  
;; PRIOR FILING DATE: 2001-03-12

;; PRIOR APPLICATION NUMBER: 09/523,095  
;; PRIOR FILING DATE: 2000-03-10  
;; PRIOR APPLICATION NUMBER: JP 2000-115246  
;; PRIOR FILING DATE: 2000-04-17  
;; PRIOR APPLICATION NUMBER: JP 2000-321822  
;; PRIOR FILING DATE: 2000-10-20  
;; NUMBER OF SEQ ID NOS: 103  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 91  
;; LENGTH: 136  
;; TYPE: PRT  
;; ORGANISM: Mus sp.  
;; FEATURE:  
;; OTHER INFORMATION: pGEM-M1H  
US-10-221-131-91

Query Match 16.7%; Score 61.5; DB 6; Length 136;  
Best Local Similarity 30.6%; Pred. No. 2.4;  
Matches 19; Conservative 9; Mismatches 21; Indels 13; Gaps 2;  
QY 20 YADGKTQTAEFKG---TFEETATBAYRYADLLAKEN-----GKTYVDVADKGYTL 66  
Db 73 YNDGTYNEKFKGKATLTSEKSSSAAYMELSSLAESDSAVYVCARGYYSYDDMGQGTTL 132  
QY 67 NI 68  
Db 133 TV 134

RESULT 9  
US-09-200-650E-3  
;; Sequence 3, Application US/09200650E  
;; GENERAL INFORMATION:  
;; APPLICANT: Patti, Joseph M.  
;; APPLICANT: Foster, Timothy J.  
;; APPLICANT: Hook, Magnus A.O.  
;; APPLICANT: Bidhinn, Deirdre Ni  
;; APPLICANT: Perkins, Samuel L.  
;; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus  
;; FILE REFERENCE: P06283US2/BAS  
;; CURRENT APPLICATION NUMBER: US/09/200,650E  
;; CURRENT FILING DATE: 1998-11-25  
;; PRIOR APPLICATION NUMBER: 60/066,815  
;; PRIOR FILING DATE: 1997-11-26  
;; PRIOR APPLICATION NUMBER: 60/098,427  
;; PRIOR FILING DATE: 1998-08-31  
;; NUMBER OF SEQ ID NOS: 23  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 3  
;; LENGTH: 930  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-09-200-650E-3

Query Match 16.5%; Score 61; DB 5; Length 930;  
Best Local Similarity 31.4%; Pred. No. 30;  
Matches 22; Conservative 16; Mismatches 24; Indels 8; Gaps 5;  
QY 4 TPEEPK-BEVTIKANLIYA-DGKTOTAB-FKG---TFEETATBAYRYADLLAKENGK--Y 55  
Db 407 TPDTSKLDKDVTDQDFVIYSDNKNKTATVLMKGTSSNKQYIIQQVAYPONSSTYDNGKIDY 466  
QY 56 TVDVADKGYT 65  
Db 467 TLDTDKTKYS 476

RESULT 10  
US-09-200-650E-7  
;; Sequence 7, Application US/09200650E  
;; GENERAL INFORMATION:  
;; APPLICANT: Patti, Joseph M.  
;; APPLICANT: Foster, Timothy J.

```

: APPLICANT: Hook, Magnus A.O.
: APPLICANT: Eichinn, Deirdre Ni
: APPLICANT: Perkins, Samuel L.
: TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
: FILE REFERENCE: P06283US2/BAS
: CURRENT APPLICATION NUMBER: US/09/200,650E
: CURRENT FILING DATE: 1998-11-25
: PRIOR APPLICATION NUMBER: 60/066,815
: PRIOR FILING DATE: 1997-11-26
: PRIOR APPLICATION NUMBER: 60/098,427
: PRIOR FILING DATE: 1998-08-31
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 7
: LENGTH: 1166
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
: US-09-200-650E-7

```

```

Query Match      16.5%; Score 61; DB 5; Length 1166;
Best Local Similarity 34.0%; Pred. No. 40;
Matches 16; Conservative 7; Mismatches 20; Indels 4; Gaps 2;

Qy      24  K T O A T F K G T F E E T A B A Y R A D L L A K E N G Y T V D V - A D K G Y T L N I K   69
          : : : : : : : : : : : : : : : : : :
Db      752  K D S T G K V I G T T D A S G K Y K E T D L - - - D N G N Y T F E F E T P A G Y T P T V K   795

```

RESULT 11  
US-10-613-520-1638  
; Sequence 1638, Application US/10613520  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; APPLICANT: Wu, Wei  
; APPLICANT: Maolong, Lu  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)E  
; CURRENT APPLICATION NUMBER: US/10/613,520  
; CURRENT FILING DATE: 2003-07-02  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 10/389,566  
; PRIOR FILING DATE: 2003-03-14  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2578  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1638  
; LENGTH: 636  
; TYPE: PRT  
; ORGANISM: Versinia pestis  
US-10-613-520-1638

Query Match	16.4%	Score 60.5;	DB 6;	Length 636;
Best Local Similarity	32.0%;	Pred. No. 21;		
Matches	24;	Conservative	4; Mismatches	32; Indels 15; Gaps 3;
Qy	2	EKTPE----	BPKEVTIKANLIYADGKTYQT-----EFKGTPEATAEAYRYADLL	48
Dd	124	KKTAEDYLGEFVTEAVITVPAYFNDAQRTKDAGRAGLEVKRIINEPTAALAYG--L	181	
Qy	49	AKENGKYTVDDVADKG	63	
Dd	182	DKEVGNETIAYVDLG	196	

RESULT 12  
US-09-897-516A-7157  
; Sequence 7157, Application US/09897516A  
; GENERAL INFORMATION:

[illegible]

QY 49 AKENGKYTVDVADKG 63  
|||:|  
Dp 182 DKEVGNRTIAVYDLG 196

RESULT 13  
US-10-613-520-1341  
Sequence 1341, Application US/10613520  
GENERAL INFORMATION:  
APPLICANT: Monsanto Technology, LLC  
APPLICANT: Laurie, Cathy C  
APPLICANT: Wu, Wei  
APPLICANT: Maolong, Lu  
TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
FILE REFERENCE: 38-77(52900)E  
CURRENT APPLICATION NUMBER: US/10/613,520  
CURRENT FILING DATE: 2003-07-02  
PRIOR APPLICATION NUMBER: US 60/365,301  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: US 10/389,566  
PRIOR FILING DATE: 2003-03-14  
PRIOR APPLICATION NUMBER: US 60/391,786  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/392,018  
PRIOR FILING DATE: 2002-06-26  
NUMBER OF SEQ ID NOS: 2578  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1341  
LENGTH: 638  
TYPE: PRT  
ORGANISM: Salmonella enterica  
US-10-613-520-1341

[illegible]

261 BEFAVSEKNSRITIKROAPSYRETTOQNTNLAY-----TGKDLGFVEKLDANAY----- 309

Db 261 BEFAVSEKNSRITIKROAPSYRETTOQNTNLAY-----TGKDLGFVEKLDANAY----- 309  
Qy 48 LAKENGKYTVDDVADKGYTLNIK 69  
Db 310 -VLEKKRYSAADDKONGYAGNVK 330  
Search completed: September 3, 2003, 11:40:56  
Job time : 11.1096 secs

RESULT 14  
US-10-613-520-1356  
; Sequence 1356, Application US/10613520  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; APPLICANT: Wu, Wei  
; APPLICANT: Maolong, Lu  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)E  
; CURRENT APPLICATION NUMBER: US/10/613,520  
; CURRENT FILING DATE: 2003-07-02  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 10/389,566  
; PRIOR FILING DATE: 2003-03-14  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2578  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1356  
; LENGTH: 638  
; TYPE: PRT  
; ORGANISM: Salmonella typhimurium  
US-10-613-520-1356

Query Match 16.4%; Score 60.5; DB 6; Length 638;  
Best Local Similarity 32.0%; Pred. No. 21;  
Matches 24; Conservative 4; Mismatches 32; Indels 15; Gaps 3;  
Qy 2 EXTPE-----EPKEVTIRANLIYADGKCTA-----EPKGTFEATAEAYADLL 48  
Db 124 KKTAEADYLGEPTVEAVITVPYFNDQROQATKDGRIAGLEVKRIINEBPTAAALAYG--L 181  
Qy 49 AKENGKYTVDDVADKG 63  
Db 182 DREVGNRITAVYDLG 196

RESULT 15  
US-10-467-534-97  
; Sequence 97, Application US/10467534  
; GENERAL INFORMATION:  
; APPLICANT: Berthet, Francois-Xavier Jacques  
; APPLICANT: Lobet, Yves  
; APPLICANT: Poolman, Jan  
; APPLICANT: Verilant, Vincent Georges Christian Louis  
; TITLE OF INVENTION: Vaccine Composition  
; FILE REFERENCE: B45261  
; CURRENT APPLICATION NUMBER: US/10/467,534  
; CURRENT FILING DATE: 2003-08-08  
; PRIOR APPLICATION NUMBER: PCT/EP02/01356  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: GB 0103169.9  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 97  
; LENGTH: 713  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-534-97

Query Match 16.4%; Score 60.5; DB 6; Length 713;  
Best Local Similarity 26.8%; Pred. No. 25;  
Matches 22; Conservative 7; Mismatches 28; Indels 25; Gaps 3;  
Qy 1 KEKTPPEPKKEVTIK-----ANLIYADGKTOTAEFKGTFEATAEAYADL 47



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:04:52 ; Search time 9.12329 Seconds  
(without alignments)  
371.130 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_153\_224

Perfect score: 369

Sequence: 1 KEKTEPEKPEVTKANLIY.....GKYTVDVADKGYTLNKFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.5	24.0	398	1	RPC4 HUMAN
2	83.5	22.6	398	1	RPC4 MOUSE
3	67	18.2	953	1	LKAI PASHA
4	67	18.2	953	1	LKAB PASHA
5	65.5	17.8	471	1	TNAI ECO57
6	65.5	17.8	471	1	TNAI ECO57
7	65.5	17.8	1179	1	DE3A STROCO
8	65	17.6	367	1	SYA AQUUPY
9	64.5	17.5	367	1	GPR OCEIH
10	64	17.3	621	1	LAMC DROME
11	64	17.3	634	1	GYRB BORHE
12	63.5	17.2	425	1	IRK6 MOUSE
13	63.5	17.2	425	1	IRK6 RAT
14	63.5	17.2	1461	1	ALOB HUMAN
15	63	17.1	6632	1	UNB9 CABEL
16	62.5	16.9	263	1	YCY3 ASTLO
17	62.5	16.9	423	1	IRK6 HUMAN
18	62.5	16.9	425	1	IRK6 MESAU
19	61.5	16.7	300	1	SIAL PIG
20	61.5	16.7	451	1	Y996 METJA
21	61.5	16.7	550	1	THS FYRAB
22	61.5	16.7	892	1	IF2 YERPE
23	61	16.5	420	1	ASSY YEAST
24	60.5	16.4	238	1	RNPH VIBCH
25	60.5	16.4	269	1	T2S1 STRFI
26	60.5	16.4	345	1	QURA LACLA
27	60.5	16.4	444	1	SLAP LACAC
28	60.5	16.4	636	1	DNK1 YERPE
29	60.5	16.4	637	1	DNK1 SALT1
30	60.5	16.4	637	1	DNK1 SALT1
31	60.5	16.4	1023	1	HLV1 ECOLI
32	60	16.3	436	1	RBL EUGST
33	60	16.3	1131	1	PMAL1 DUNBI

34	59.5	16.1	373	1	CONS ARATH
35	59.5	16.1	390	1	IRKB_CAVPO
36	59.5	16.1	796	1	YEZ9 YEAST
37	59.5	16.1	1454	1	CSP2 HUMAN
38	59.5	16.1	1857	1	FAS2 PENPA
39	59	16.0	322	1	MRPH METKA
40	59	16.0	366	1	RUVB ANASP
41	59	16.0	776	1	RTNI HUMAN
42	59	16.0	1829	1	Y296 HUMAN
43	58.5	15.9	417	1	PUR2 THETN
44	58.5	15.9	424	1	PUR2 AQUAE
45	58.5	15.9	1001	1	ORK1 DROME

## ALIGNMENTS

RESULT 1  
RPC4\_HUMAN STANDARD; PRT; 398 AA.  
AC P05423; Q9BPV7; Q9BP21; Q9BXB3;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE DNA-directed RNA polymerase III 47 kDa polypeptide (EC 2.7.7.6) (RNA  
DE polymerase C subunit 4) (RPC4) (RPC53) (BNS1 protein).  
GN ENSIT OR BNS1 OR POLR3D.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88065472; PubMed=3683386;  
RA Ittmann M., Greco A., Basilico C.;  
RT "Isolation of the human gene that complements a temperature-sensitive  
RT cell cycle mutation in BHK cells.";  
RL Mol. Cell. Biol. 7:3386-3393(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22278866; PubMed=12391170;  
RA Hu P., Wu S., Sun Y., Yuan C.-C., Kobayashi R., Myers M.P.,  
RA Hernandez N.;  
RT "Characterization of human RNA polymerase III identifies orthologues  
RT for Saccharomyces cerevisiae RNA polymerase III subunits.";  
RL Mol. Cell. Biol. 22:8044-8055(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung, and Lymph;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
RA Schnerch A., Schin J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 1-50 FROM N.A.  
RX MEDLINE=21282972; PubMed=11279001;

Chong S.S., Hu P., Hernandez N.;  
 "Reconstitution of transcription from the human U6 small nuclear RNA  
 promoter with eight recombinant polypeptides and a partially purified  
 RNA polymerase III complex.";  
 J. Biol. Chem. 276:20727-20734(2001).  
 [5]  
 RN SEQUENCE OF 1-12 FROM N.A.  
 RP Ittmann M.;  
 RA "Promoter structure and cell cycle control of the RN51 cell cycle  
 gene, which encodes a subunit of RNA polymerase III.";  
 RT Submitted (Oct-1993) to the EMBL/GenBank/DBJ databases.  
 RL  
 CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
 of DNA into RNA using the four ribonucleoside triphosphates as  
 CC substrates. Complements a temperature-sensitive cell cycle  
 CC mutation in BHK cells.  
 CC  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC  
 CC -!- SUBUNIT: RNA polymerase III consists of about 15 different  
 CC subunits. Interacts with RPC5.  
 CC  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC  
 CC -!- CAUTION: Ref.3 (AAH03039 and AAH00516) sequence differs from that  
 CC shown due to a frameshift in position 175.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M17754; AAA51838.1; -;  
 DR EMBL; AY092086; AAM18216.1; -;  
 DR EMBL; BC002603; AAH02603.1; -;  
 DR EMBL; BC003039; AAH03039.1; ALT\_FRAME.  
 DR EMBL; BC004484; AAH04484.1; -;  
 DR EMBL; BC000516; AAH00516.1; ALT\_FRAME.  
 DR EMBL; AF346574; AAK15371.1; -;  
 DR EMBL; L15301; AA72377.1; -;  
 DR PIR; A43700; A43700.  
 DR Genew; HGNC:1080; BNS1T.  
 DR MIM; 187280; -;  
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.  
 DR Pfam; PF05132; RNA pol Rpc4; 1.  
 DR Transfaser; DNA-directed RNA polymerase; Transcription;  
 KW Nuclear protein.  
 FT CONFLICT 9 18 EBPSTGGPRP -> RPARGPDL (IN REF. 1).  
 FT EPTST -> RPAR (IN REF. 5).  
 FT CONFLICT 9 12 LIGRRPAPLTPGRL -> SSGGGGLPSPAV (IN REF.  
 FT CONFLICT 26 40 1).  
 FT  
 FT CONFLICT 98 98 G -> R (IN REF. 1).  
 FT CONFLICT 233 233 K -> R (IN REF. 2).  
 FT CONFLICT 286 286 P -> L (IN REF. 2).  
 SQ SEQUENCE 398 AA; 44396 MW; CDSAFF3257B78410 CRC64;  
 Query Match 24.0%; Score 88.5; DB 1; Length 398;  
 Best Local Similarity 33.3%; Pred. No. 0.052;  
 Matches 27; Conservative 9; Mismatches 26; Indels 19; Gaps 3;  
 QY 3 KTPBEPKERTVKANLIYAD-----GKTQTAEFGKTFEATAEAYRYDLAK 50  
 Db 66 KIKEPKEVTVKKEKRDRDRQREGHGRGRPEVQSHIFEGGPAEMMK-----KK 120  
 QY 51 ENKQTVDVADKG--YTLNIK 69  
 Db 121 GNWDKTVDSVDMGPHSHIINIK 141  
 RESULT 2  
 RPC4 MOUSE  
 ID RPC4\_MOUSE STANDARD; PRT; 398 AA.  
 AC Q91WMI; Q9CZ02;  
 DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE DNA-directed RNA polymerase III 47 kDa polypeptide (EC 2.7.7.6) (RNA  
 DE polymerase C subunit 4) (RPC4).  
 GN ENS1T OR POLR3D.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 CC SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Mateu V., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection";  
 RT Nature 409:685-690(2001).  
 RL  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Salivary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
 CC of DNA into RNA using the four ribonucleoside triphosphates as  
 CC substrates (By similarity).  
 CC  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC  
 CC -!- SUBUNIT: RNA polymerase III consists of about 15 different  
 CC subunits. Interacts with RPC5 (By similarity).  
 CC  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

```

CC -----
DR EMBL; AK013165; BAB28687.1; -.
DR EMBL; BC016102; AAH16102.1; -.
DR MGD; MGI:1914315; 2810426M17R1k.
DR Pfam; PF05132; RNA_pol_Rpc4; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription;
KW Nuclear protein.
FT CONFLICT 201 201 L -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 44323 MW; D509DE632C656FFD CRC64;

Query Match
Best Local Similarity 22.6%; Score 83.5; DB 1; Length 398;
Matches 26; Conservative 10; Mismatches 26; Indels 19; Gaps 3;

QY 3 KTPPEPKKEVTKANLIYAD-----GKTQTAEFKGTPEATAEAYRYADLLAK 50
Db 66 KIKPEPKKEVTKKKEKRRDRQRQHGGRGRGRDEVIGSHSIFEQGPAEMK-----KK 120
QY 51 ENKQYTVADKG--YTLNIK 69
Db 121 GNWDKTVDSMDGSPSHIINIK 141

RESULT 3
LKAB_PASHA
ID LKAB_PASHA STANDARD; PRT; 953 AA.
AC P16535;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukotoxin from serotype A1.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=87306837; PubMed=3040588;
RA Lo R.Y.C., Strathee C.A., Shewen P.B.;
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella
haemolytica A1.";
RL Infect. Immun. 55:1987-1996(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1 / PHL101;
RX MEDLINE=89210283; PubMed=2707120;
RA Highlander S.K., Chidambaram M., Engler M.J., Weinstein G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
cluster.";
RN DNA 8:15-28(1989).
RN [3]
RP SEQUENCE OF 884-953 FROM N.A.
RC STRAIN=Serotype A1 / PHL101;
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstein G.M.;
RT "Secretion and expression of the Pasteurella haemolytica Leukotoxin.";
RL J. Bacteriol. 172:2343-2350(1990).
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture by mechanisms not clearly
defined.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
calcium, which is required for target cell-binding or cytolytic
activity.
CC -!- DOMAIN: The three transmembrane domains are believed to be
involved in pore formation by the cytotoxin (By similarity).
CC -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
modified (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----

```

----- This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20730; AAA25529.1; -.
DR EMBL; M24197; AAA25543.1; -.
DR PIR; B30169; B30169.
DR InterPro; IPR001343; Hemlyan_Ca_bind.
DR InterPro; IPR003995; RtxA.
DR Pfam; PF00353; hemolysinCbind; 5.
DR Pfam; PF02382; RTX; 1.
DR PRINTS; PR00313; CAENDNGRPT.
DR PRINTS; PR01488; RTXTOXINA.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
KW Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW Lipoprotein; Palmitate.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 297 318 POTENTIAL.
FT TRANSMEM 366 390 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
FT CONFLICT 409 414 FEHVAN -> LSTLOI (IN REF. 2).
FT CONFLICT 742 742 D -> Y (IN REF. 2).
SQ SEQUENCE 953 AA; 101996 MW; 7F93D113A118C05F CRC64;

Query Match
Best Local Similarity 18.2%; Score 67; DB 1; Length 953;
Matches 21; Conservative 8; Mismatches 25; Indels 10; Gaps 1;

QY 9 KKEVTKANLIYADKGTQTAEFKGTPEATAE-----AYRYADLLAKENGYTV 58
Db 832 KEKVTIQNWFEADFAKEVPVNYKATKDEIIGQNGERITSQVDDLIAGNGKITOD 891
QY 59 VADK 62
Db 892 ELKS 895

RESULT 4
LKAB_PASHA
ID LKAB_PASHA STANDARD; PRT; 953 AA.
AC P55118;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukotoxin from serotype A1.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype A1;
RX MEDLINE=94041617; PubMed=8225575;
RA Burrows L.L., Olah-Winfield E., Lo R.Y.C.;
RT "Molecular analysis of the leukotoxin determinants from Pasteurella
haemolytica serotypes 1 to 16.";
RL Infect. Immun. 61:5001-5007(1993).
CC -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
cell membranes and cause cell rupture by mechanisms not clearly
defined.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The Gly-rich region is probably involved in binding
calcium, which is required for target cell-binding or cytolytic
activity.

```

CC activity.

CC -!- DOMAIN: The three transmembrane domains are believed to be

CC involved in pore formation by the cytotoxin (BY SIMILARITY).

CC -!- PFM: Palmitoylated by lktC. The toxin only becomes active when

CC modified (by similarity).

CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U01215; AAB36689.1; --

CC InterPro; IPR001343; Hemlysn\_Ca\_bind.

CC InterPro; IPR003995; RtxA.

CC Pfam; PF00353; hemolysinCbind; 5.

CC Pfam; PF02382; RTX; 1.

CC PRINTS; PR00313; CABNDNGRPT.

CC PRINTS; PR01488; RTXTOXINA.

CC PROSITE; PS00330; HEMOLYSIN\_CALCIIUM; 4.

CC Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;

CC Lipoprotein; Palmitate.

CC TRANSMEM 230 250 POTENTIAL.

CC TRANSMEM 297 317 POTENTIAL.

CC TRANSMEM 381 401 POTENTIAL.

CC DOMAIN 734 784 6 X REPEATS, GLY-RICH.

CC REPEAT 734 739 1.

CC REPEAT 743 748 2.

CC REPEAT 752 757 3.

CC REPEAT 761 766 4.

CC REPEAT 770 775 5.

CC REPEAT 779 784 6.

CC SEQUENCE 953 AA; 102206 MW; 927FF56CFC884F12 CRC64;

CC -----

CC Query Match 18.2%; Score 67; DB 1; Length 953;

CC Best Local Similarity 32.8%; Pred. No. 23;

CC Matches 21; Conservative 8; Mismatches 25; Indels 10; Gaps 1;

CC -----

CC QY 9 KEVTKANLIYADGKTQTAEPKGTPEATAE-----AYRYADLLAKENGKYYTD 58

CC Db 832 KEKVTIQDFREADFAKEVRNYKATDKIEELIGQNGRITSKQVDDLLAKNGKITQD 891

CC QY 59 VADK 62

CC Db 892 ELSK 895

CC -----

CC RESULT 5

CC TNA\*\_ECOS7 STANDARD; PRT; 471 AA.

CC ID TNA\*\_ECOS7 STANDARD; PRT; 471 AA.

CC AC O8XB34;

CC DT 28-FEB-2003 (Rel. 41, Created)

CC DT 28-FEB-2003 (Rel. 41, Last sequence update)

CC DT 28-FEB-2003 (Rel. 41, Last annotation update)

CC DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).

CC GN TNA\*\_OR Z503 OR ECS4645.

CC OS Escherichia coli O157:H7.

CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

CC OC Enterobacteriaceae; Escherichia.

CC OX NCBI\_TaxID=83334;

CC RN [1]

CC SEQUENCE FROM N.A.

CC STRAIN=O157:H7 / EDL933 / ATCC 700927;

CC RX MEDLINE=21074935; PubMed=11206551;

CC RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

CC Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

CC Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

CC Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

CC Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;"

RL Nature 409:529-533 (2001).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=O157:H7 / RIMD 0509952;

CC MEDLINE=21156231; PubMed=11258796;

CC Hayashi T., Makino K., Ohnishi M., Kurokawa K., Yokoyama K.,

CC Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,

CC Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,

CC Kuhara S., Shiba T., Hattori M., Shinagawa H.;

CC "Complete genome sequence of enterohaemorrhagic Escherichia coli

CC O157:H7 and genetic comparison with a laboratory strain K-12.;"

CC DNA Res. 8:11-22 (2001).

CC -!- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +

CC NH(3)

CC -!- COFACTOR: Pyridoxal phosphate (By similarity).

CC -!- PATHWAY: Tryptophan catabolism.

CC -!- SUBUNIT: Homotetramer (By similarity).

CC -!- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AB005602; AAG58908.1; ALT\_INIT.

CC EMBL; AP002566; BAB38068.1; ALT\_INIT.

CC HAMAP; MF\_00544; -; 1.

CC InterPro; IPR001597; Beta\_elim\_lyase.

CC Pfam; PF01212; Beta\_elim\_lyase; 1.

CC PRODOM; PD005927; Beta\_elim\_lyase; 1.

CC PROSITE; PS00853; BETA\_ELIM\_LYASE; 1.

CC KW Tryptophan catabolism; lyase; Pyridoxal phosphate; Complete proteome.

CC FT BINDING 270 270 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC SQ SEQUENCE 471 AA; 52789 MW; 49DD2D41BD9D0034 CRC64;

CC -----

CC Query Match 17.8%; Score 65.5; DB 1; Length 471;

CC Best Local Similarity 51.9%; Pred. No. 16;

CC Matches 14; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

CC -----

CC QY 24 KTQTAEPKGT-TFEATAEAYRYADLLA 49

CC Db 239 KQREAYKDWITIEQITRETYKYADMLA 265

CC -----

CC RESULT 6

CC TNA\*\_ECOLI STANDARD; PRT; 471 AA.

CC ID TNA\*\_ECOLI STANDARD; PRT; 471 AA.

CC AC P00913; P78123;

CC DT 21-JUL-1986 (Rel. 01, Created)

CC DT 01-JUL-1993 (Rel. 26, Last sequence update)

CC DT 28-FEB-2003 (Rel. 41, Last annotation update)

CC DE Tryptophanase (EC 4.1.99.1) (L-tryptophan indole-lyase) (TNase).

CC GN TNA\*\_OR IND OR B3708 OR SF3754.

CC OS Escherichia coli, and

CC OS Shigella flexneri.

CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

CC OC Enterobacteriaceae; Escherichia.

CC OX NCBI\_TaxID=562, 623;

CC RN [1]

CC SEQUENCE FROM N.A.

CC SPECIES=E.coli; STRAIN=K12;

CC RX MEDLINE=82007678; PubMed=6268608;

CC RA Deeley M.C., Yanofsky C.;

CC "Nucleotide sequence of the structural gene for tryptophanase of

CC Escherichia coli K-12.;"

CC RL J. Bacteriol. 147:787-796 (1981).

CC [2]

CC SEQUENCE FROM N.A., AND IMPORTANCE OF CYS-298.

CC SPECIES=E.coli; STRAIN=B/1t7-A;



RX MEDLINE=89323226; PubMed=2502187;  
 RA Tokunaga M., Tsujimoto N., Oda T., Honda T., Yumoto N., Ito S.,  
 RA Yamamoto M., Kim E.H., Hiragi Y.,  
 RT "Role of cysteine residues in tryptophanase for monovalent cation-  
 RT induced activation.";  
 RL Biochimie 71:711-720(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12 / MGL655;  
 RX MEDLINE=93315143; PubMed=766882;  
 RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner P.R.,  
 RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli  
 RT genome: organizational symmetry around the origin of replication.";  
 RL Genomics 16:551-561(1993).  
 RN [4]  
 RP SEQUENCE OF TRYPTIC PEPTIDES.  
 RC SPECIES=E.coli; STRAIN=K12;  
 RX MEDLINE=72134434; PubMed=4551944;  
 RA Kagamiyama H., Matsubara H., Snell E.E.,  
 RT "The chemical structure of tryptophanase from Escherichia coli. 3.  
 RT Isolation and amino acid sequence of the tryptic peptides.";  
 RL J. Biol. Chem. 247:1576-1586(1972).  
 RN [5]  
 RP SEQUENCE OF 1-21 FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12;  
 RX MEDLINE=86033634; PubMed=3902796;  
 RA Stewart V., Yanofsky C.,  
 RT "Evidence for transcription antitermination control of tryptophanase  
 RT operon expression in Escherichia coli K-12.";  
 RL J. Bacteriol. 164:731-740(1985).  
 RN [6]  
 RP SEQUENCE OF 463-471 FROM N.A.  
 RC SPECIES=E.coli;  
 RX MEDLINE=91216998; PubMed=2022620;  
 RA Sarsero J.P., Wooley P.J., Gollnick P.D., Yanofsky C., Pittard A.J.,  
 RT "A new family of integral membrane proteins involved in transport of  
 RT aromatic amino acids in Escherichia coli.";  
 RL J. Bacteriol. 173:3231-3234(1991).  
 RN [7]  
 RP SEQUENCE OF 1-12.  
 RC SPECIES=E.coli; STRAIN=K12 / EMG2;  
 RX MEDLINE=97443975; PubMed=9298646;  
 RA Link A.J., Robison K., Church G.M.,  
 RT "Comparing the predicted and observed properties of proteins encoded  
 RT in the genome of Escherichia coli K-12.";  
 RL Electrophoresis 18:1259-1313(1997).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
 RX MEDLINE=2272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang P., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.,  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 RT through comparison with genomes of Escherichia coli K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 RN [9]  
 RP MUTAGENESIS OF CYS-294 AND CYS-298.  
 RC SPECIES=E.coli;  
 RX MEDLINE=89278130; PubMed=2659590;  
 RA Phillips R.S., Gollnick P.D.,  
 RT "Evidence that cysteine 298 is in the active site of tryptophan  
 RT indole-lyase.";  
 RL J. Biol. Chem. 264:10627-10632(1989).  
 CC -I- COFACTOR: Pyridoxal phosphate.  
 CC -I- PATHWAY: Tryptophan catabolism.  
 CC -I- SUBUNIT: Homotetramer.  
 CC -I- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; K00032; AAA24676.1; -;  
 DR EMBL; X15974; CAA34096.1; -;  
 DR EMBL; M11990; AAA24679.1; -;  
 DR EMBL; M59914; -; NOT ANNOTATED CDS.  
 DR EMBL; L10328; AAA62059.1; ALT\_INIT.  
 DR EMBL; AE000448; AAC76731.1; ALT\_INIT.  
 DR EMBL; AE015382; AAN45197.1; ALT\_INIT.  
 DR PIR; B65173; WZEC.  
 DR HSSP; P28796; 1AX4.  
 DR ECODBASE; G046.5; 6TH EDITION.  
 DR HAMAP; MF\_00544; -; 1.  
 DR InterPro; IPR001597; Beta\_elim\_lyase.  
 DR Pfam; PF01212; Beta\_elim\_lyase; 1.  
 DR ProDom; PD005927; Beta\_elim\_lyase; 1.  
 DR PROSITE; PS00853; BETA\_ELIM\_LYASE; 1.  
 KW Tryptophan catabolism; Lyase; Pyridoxal phosphate; Complete proteome.  
 FT BINDING 270 270 PYRIDOXAL PHOSPHATE.  
 FT MUTAGEN 294 294 C->S: IDENTICAL TO WILD-TYPE.  
 FT CONFLICT 137 140 C->S: ALTERED ACTIVITY.  
 FT CONFLICT 379 380 DTTO -> TTQG (IN REF. 1).  
 FT CONFLICT 379 380 QA -> TG (IN REF. 1 AND 2).  
 SQ SEQUENCE 471 AA; 52773 MW; 5APC1F41BD9D0034 CRC64;  
 Query Match 17.8%; Score 65.5; DB 1; Length 471;  
 Best Local Similarity 51.9%; Pred. No. 16;  
 Matches 14; Conservative 5; Mismatches 7; Indels 1; Gaps 1;  
 QY 24 KTQTAEFKG-TPEATAEAYRYADLLA 49  
 DB 239 KQEAAYKDWTEIQITRETYKADMLA 265  
 RESULT 7  
 ID DP3A STRCO STANDARD; PRT; 1179 AA.  
 AC Q9Z618;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE DNA polymerase III alpha subunit (EC 2.7.7.7).  
 GN DNAE OR SCO2064 OR SC4G6.33C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxId=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=99157597; PubMed=10048037;  
 RA Flett E., Jungmann-Campello D., Mersinias V., Koh S.L.-M., Godden R.,  
 RA Smith C.P.,  
 RT "A 'Gram-negative-type' DNA polymerase III is essential for  
 RT replication of the linear chromosome of Streptomyces coelicolor  
 RT A3(2)."  
 RL Mol. Microbiol. 31:949-958(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kisser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kisser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RT Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME  
 CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.  
 CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.  
 CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 CC + [DNA] (N).  
 CC -!- SUBUNIT: DNA polymerase III contains a core (composed of alpha,  
 CC epsilon and theta chains) that associates with a tau subunit. This  
 CC core dimerizes to form the POLIII' complex. POLIII' associates  
 CC with the gamma complex (composed of gamma, delta, delta', psi and  
 CC chi chains) and with the beta chain to form the complete DNA  
 CC polymerase III complex (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. DNAE  
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF108191; AAD16978.1; -;  
 DR EMBL: AL096884; CAB51456.1; -;  
 DR PIR: T35093; T35093.  
 DR InterPro: IPR004013; PHP\_C.  
 DR InterPro: IPR003141; PHP\_N.  
 DR InterPro: IPR004805; PolC\_alpha.  
 DR InterPro: IPR004365; trna\_anti.  
 DR Pfam: PF02811; PHP\_C; 1.  
 DR Pfam: PF02231; PHP\_N; 1.  
 DR Pfam: PF01336; trna\_anti; 1.  
 DR SMART: SM00481; POLIITAC; 1.  
 DR TIGRFAMs: TIGR00594; polc; 1.  
 DR Transferrase; DNA-directed DNA polymerase; DNA replication;  
 KW Complete proteome. 802 E -> X (IN MUTANT TS-38).  
 FT VARIANT 802 802 E -> X (IN MUTANT TS-38).  
 SQ SEQUENCE 1179 AA; 130795 MW; 7E4B58675B634CD3 CRC64;

Query Match 17.8%; Score 65.5; DB 1; Length 1179;  
 Best Local Similarity 29.9%; Pred. No. 42;  
 Matches 20; Conservative 12; Mismatches 28; Indels 7; Gaps 2;

QY 2 EKTPEEPKKEVT-----IKANLIYADGKTQTAEFGTPEATAEAYRVADLLAKENGKYYT 56  
 Db 131 QKWPMDKETISQWSEGVASTGCPSEGVQTLRLGLHFDALKAAADYDIFGKD--RYF 188

QY 57 VDVAKG 63

Db 189 LELMDHG 195

## RESULT 8

SVL\_AQUYP STANDARD; PRT; 871 AA.  
 ID SVL\_AQUYP  
 AC Q9XDM3;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (Alars).  
 GN ALAS.  
 OS Aquifex pyrophilus.  
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
 CC NCBI\_TaxID=2714;

RN SEQUENCE FROM N.A.  
 RP STRAIN=K015A / DSM 6858;  
 RX MEDLINE=99214260; PubMed=10198119;  
 RA Klenk H.-P., Meier T.D., Durovic P., Schwass V., Lottepeich F.,  
 RA Dennis P.P., Zillig W.;  
 RT "RNA polymerase of Aquifex pyrophilus: implications for the evolution  
 RT of the bacterial rpoBC operon and extremely thermophilic bacteria.";  
 RL J. Mol. Evol. 48:528-541(1999).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +  
 CC diphosphate + L-alanyl-tRNA(Ala).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF027500; AAD25871.1; -;  
 DR HAMAP: MF\_00036; -; 1.  
 DR InterPro: IPR003156; DHHA1.  
 DR InterPro: IPR002318; tRNA-synt\_2c.  
 DR InterPro: IPR006193; tRNA\_synt\_Ala.  
 DR Pfam: PF02272; DHHA1; 1.  
 DR Pfam: PF01411; tRNA-synt\_2c; 1.  
 DR PRINTS: PR00980; TRNASYNTHALA.  
 DR TIGRFAMs: TIGR00344; alas; 1.  
 DR PROSITE: PS50860; AA TRNA LIGASE II ALA; 1.  
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 SQ SEQUENCE 871 AA; 99395 MW; D2D15915FC00CC31 CRC64;

Query Match 17.6%; Score 65; DB 1; Length 871;  
 Best Local Similarity 30.2%; Pred. No. 34;  
 Matches 19; Conservative 8; Mismatches 20; Indels 16; Gaps 2;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFGTPEATAEAYRVADLLAKENGKYYTVDV 59  
 Db 754 REKVKERHRELT-----YYGVFEDVEPEELRLNLIADILRQTGKDVVFI 798

QY 60 ADK 62

Db 799 ASK 801

## RESULT 9

GPR\_OCEIH STANDARD; PRT; 367 AA.  
 ID GPR\_OCEIH  
 AC Q8XC9;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Germination protease precursor (EC 3.4.99.-) (Spore protease).  
 GN GPR OR 0B1975.  
 OS Oceanobacillus thelyensis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
 CC NCBI\_TaxID=182710;

## [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=HT831 / DSM 14371 / JCM 11309;  
 RX MEDLINE=2220767; PubMed=12235376;  
 RA Takami H., Takaki Y., Uchiyama I.;  
 RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya  
 RT Ridge and its unexpected adaptive capabilities to extreme  
 RT environments";  
 RL Nucleic Acids Res. 30:3927-3935(2002).  
 CC -!- FUNCTION: Initiates the rapid degradation of small, acid-soluble  
 CC proteins during spore germination (By similarity).  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- PTM: Autoproteolytically processed. The inactive tetrameric



DR HSSP; P06982; 1AJ6.  
 DR InterPro; IPR003594; ATPbind ATPase.  
 DR InterPro; IPR002288; DNA\_gyraseB\_C.  
 DR InterPro; IPR000585; DNA\_gyraseB.  
 DR InterPro; IPR001241; DNA\_topoisomII.  
 DR InterPro; IPR006171; Toprim dom.  
 DR Pfam; PF00204; DNA\_gyraseB\_1.  
 DR Pfam; PF00986; DNA\_gyraseB\_C\_1.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF01751; Toprim; 1.  
 DR PRINTS; PR00418; TP12FAMILY.  
 DR ProDom; PD149633; DNA\_gyraseB\_C; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00433; TOP2c; 1.  
 DR TIGRFAMs; TIGR01059; GYR8; 1.  
 DR PROSITE; PS00177; TOPOISOMERASE\_II; 1.  
 DR Topoisomerase; Isomerase; ATP-binding.  
 SQ SEQUENCE 634 AA; 70848 MW; 5E58C5F32126833C CRC64;  
 Query Match 17.3%; Score 64; DB 1; Length 634;  
 Best Local Similarity 32.7%; Pred.No. 31; Mismatches 23; Indels 2; Gaps 1;  
 Matches 18; Conservative 12;  
 QY 17 NLIYADGKTQTAEFGTFFETAEAYRYADLLAKENGKTVVDADKGYT--LNLIK 69  
 Db 270 NINTREGGTHVAGFGKGLKAMSEAFRDSKISKQVPSLTLDFFKGLTAVISIK 324  
 RESULT 12  
 IRK6 MOUSE  
 ID TRK6 MOUSE STANDARD; PRT; 425 AA.  
 AC P48542; O70290; P70216; P70306; P70308; P70309; P70454;  
 AC Q9QVH5;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE G protein-activated inward rectifier potassium channel 2 (GIRK2)  
 DE (Potassium channel, inwardly rectifying, subfamily J, member 6)  
 DE (Inward rectifier K+ channel Kir3.2).  
 DE KCNJ6 OR KCNJ7 OR GIRK2 OR W.  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]\_\_TaxID=10090;  
 RP SEQUENCE FROM N.A. (ISOFORM GIRK2-1).  
 RC TISSUE=Brain;  
 RX MEDLINE=95010760; PubMed=7926018;  
 RA Lesage F., Duprat F., Fink M., Guillemare E., Coppola T.,  
 RA Lazdunski M., Hugnot J.-P.;  
 RT "Cloning provides evidence for a family of inward rectifier and G-  
 protein coupled K+ channels in the brain."  
 RL FEBS Lett. 353:37-42(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM GIRK2A).  
 RC TISSUE=Brain;  
 RX MEDLINE=96081927; PubMed=7499385;  
 RA Lesage F., Guillemare E., Fink M., Duprat F., Heurteaux C.,  
 RA Fosset M., Romey G., Barhanin J., Lazdunski M.;  
 RT "Molecular properties of neuronal G-protein-activated inwardly  
 rectifying K+ channels."  
 RL J. Biol. Chem. 270:28660-28667(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM GIRK2B).  
 RC TISSUE=Brain;  
 RX MEDLINE=96136315; PubMed=8573147;  
 RA Isomoto S., Kondo C., Takahashi N., Matsumoto S., Yamada M.,  
 RA Takumi T., Horio Y., Kurachi Y.;  
 RT "A novel ubiquitously distributed isoform of GIRK2 (GIRK2B) enhances  
 GIRK1 expression of the G-protein-gated K+ current in Xenopus  
 oocytes."  
 RL Biochem. Biophys. Res. Commun. 218:286-291(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN=129/Svj;  
 RX MEDLINE=98389623; PubMed=9721208;  
 RA Wei J., Hodes M.E., Piva R., Feng Y., Wang Y., Ghetti B., Dlouhy S.R.;  
 RT "Characterization of murine Kir2.2 transcript isoforms: structure and  
 differential expression."  
 RL Genomics 51:379-390(1998).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM GIRK2D).  
 RA Inanobe A., Horio Y., Fujita A., Tanemoto M., Kurachi Y.;  
 RT "Molecular cloning and characterization of a novel splicing variant of  
 Kir2.2/GIRK2 predominantly expressed in mouse testis."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP VARIANT WEAVER.  
 RX MEDLINE=96024646; PubMed=7550338;  
 RA Patil N., Cox D.R., Bhat D., Faham M., Myers R.M., Peterson A.S.;  
 RT "A potassium channel mutation in weaver mice implicates membrane  
 excitability in granule cell differentiation."  
 RL Nat. Genet. 11:126-129(1995).  
 CC -!- FUNCTION: THIS POTASSIUM CHANNEL IS CONTROLLED BY G PROTEINS. IT  
 CC PLAYS A ROLE IN GRANULE CELL DIFFERENTIATION, POSSIBLY VIA  
 CC MEMBRANE HYPERPOLARIZATION. INWARD RECTIFIER K+ CHANNELS ARE  
 CC CHARACTERIZED BY A GREATER TENDENCY TO ALLOW POTASSIUM TO FLOW  
 CC INTO THE CELL RATHER THAN OUT OF IT. THEIR VOLTAGE DEPENDENCE IS  
 CC REGULATED BY THE CONCENTRATION OF EXTRACELLULAR POTASSIUM; AS  
 CC EXTERNAL K+ IS RAISED, THE VOLTAGE RANGE OF THE CHANNEL OPENING  
 CC SHIFTS TO MORE POSITIVE VOLTAGES. THE INWARD RECTIFICATION IS  
 CC MAINLY DUE TO THE BLOCKAGE OF OUTWARD CURRENT BY INTERNAL  
 CC MAGNESIUM.  
 CC -!- SUBUNIT: MAY ASSOCIATE WITH GIRK1 OR GIRK4 TO FORM A G-PROTEIN-  
 CC ACTIVATED HETEROMULTIMER PORE-FORMING UNIT. THE RESULTING INWARD  
 CC CURRENT IS MUCH LARGER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC Event=Alternative splicing; Named isoforms=5;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=GIRK2A;  
 CC IsoId=P48542-1; Sequence=Displayed;  
 CC Name=GIRK2-1;  
 CC IsoId=P48542-2; Sequence=VSP\_002803;  
 CC Name=GIRK2B;  
 CC IsoId=P48542-3; Sequence=VSP\_002804, VSP\_002805;  
 CC Name=GIRK2C;  
 CC IsoId=P48542-4; Sequence=VSP\_002806, VSP\_002807, VSP\_002808;  
 CC Name=GIRK2D; Synonyms=Kir3.2D;  
 CC IsoId=P48542-5; Sequence=Not described;  
 CC -!- TISSUE SPECIFICITY: CEREBELLUM, TESTES, CORTEX, AND SUBSTENTIA  
 CC NIGRA.  
 CC -!- DISEASE: DEFECTS IN KCNJ6 ARE THE CAUSE OF WEAVER (WV). HOMOZYGOUS  
 CC ANIMALS SUFFER FROM SEVERE ATAXIA THAT IS OBVIOUS BY ABOUT THE  
 CC SECOND POSTNATAL WEEK. THE CEREBELLUM OF THESE ANIMALS IS  
 CC DRASTICALLY REDUCED IN SIZE DUE TO DEPLETION OF THE MAJOR CELL  
 CC TYPE OF CEREBELLUM, THE GRANULE CELL NEURON. HETEROZYGOUS ANIMALS  
 CC ARE NOT ATAXIC BUT HAVE AN INTERMEDIATE NUMBER OF SURVIVING  
 CC GRANULE CELLS. MALE HOMOZYGOTES ARE STERILE, BECAUSE OF COMPLETE  
 CC FAILURE OF SPERM PRODUCTION. BOTH HETERO- AND HOMOZYGOUS ANIMALS  
 CC UNDERGO SPORADIC TONIC-CLONIC SEIZURES.  
 CC -!- SIMILARITY: BELONGS TO THE INWARD RECTIFIER-TYPE K+ CHANNEL  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).  
 CC -----  
 CC EMBL; U37253; AAA91457.1; -;  
 CC EMBL; U11859; AAA53245.1; -;  
 CC EMBL; U51122; AAC34141.1; -;  
 CC EMBL; U51123; AAC34142.1; -;

DR EMBL; U51124; AAC34143.1; --  
 DR EMBL; U51125; AAC34144.1; --  
 DR EMBL; U51126; AAC34145.1; --  
 DR EMBL; AF040049; AAC34286.1; --  
 DR EMBL; AF040047; AAC34286.1; JOINED.  
 DR EMBL; AF040050; AAC34287.1; --  
 DR EMBL; AF040049; AAC34287.1; JOINED.  
 DR EMBL; AF040051; AAC34285.1; --  
 DR EMBL; AF040047; AAC34285.1; JOINED.  
 DR EMBL; AF040049; AAC34285.1; JOINED.  
 DR EMBL; AF040052; AAC34284.1; --  
 DR EMBL; AF040047; AAC34284.1; JOINED.  
 DR EMBL; AF040049; AAC34284.1; JOINED.  
 DR EMBL; AF040051; AAC34284.1; JOINED.  
 DR EMBL; D86040; BAA12972.1; --  
 DR EMBL; AB029502; BAA88430.1; --  
 DR PIR; S48738; S48738.  
 DR MGD; MGI:104781; Kcnj6.  
 DR GO; GO:0013467; F1G-protein activated inward rectifier potass. . . ; IDA.  
 DR InterPro; IPR001838; K-channel\_IR.  
 DR Pfam; PF01007; IRK; 1.  
 DR PRINTS; PR01320; KIRCHANNEL.  
 DR ProDom; PD001103; K-channel\_IR; 2.  
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;  
 KW Alternative splicing; Disease mutation; Potassium transport.  
 FT DOMAIN 1 96 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 97 119 M1 (POTENTIAL).  
 FT DOMAIN 144 160 H5 (POT-FORMING) (POTENTIAL).  
 FT TRANSMEM 169 193 M2 (POTENTIAL).  
 FT DOMAIN 194 425 CYTOPLASMIC (POTENTIAL).  
 FT SITE 184 184 ROLE IN THE CONTROL OF POLYAMINE-MEDIATED CHANNEL GATING AND IN THE BLOCKING BY INTRACELLULAR MAGNESIUM (BY SIMILARITY).  
 FT VARSPLIC 415 425 Missing (in isoform GIRK2-1).  
 FT VARSPLIC 319 327 /FTID=VSP\_002803.  
 FT VARSPLIC 328 425 /FTID=VSP\_002804.  
 FT VARSPLIC 1 18 Missing (in isoform GIRK2B).  
 FT VARSPLIC 319 320 Missing (in isoform GIRK2C).  
 FT VARSPLIC 321 425 MT -> QF (in isoform GIRK2C).  
 FT VARSPLIC 156 186 Missing (in isoform GIRK2C).  
 FT VARSPLIC 313 313 G -> S (IN WEAVER).  
 FT VARSPLIC 344 344 I -> M.  
 FT VARSPLIC 67 67 M -> L.  
 FT CONFLICT 260 260 S -> C (IN REF. 3 AND 4).  
 FT CONFLICT 381 381 S -> T (IN REF. 5).  
 FT CONFLICT 381 381 V -> L (IN REF. 5).  
 SQ SEQUENCE 425 AA; 48651 MW; 285153DCB1B60331 CRC64;  
 Query Match 17.2%; Score 63.5; DB 1; Length 425;  
 Best Local Similarity 31.1%; Pred. No. 23;  
 Matches 19; Conservative 12; Mismatches 23; Indels 7; Gaps 3;  
 QY 1 KEKTPPEPKVEYIKANLYAGKQTQTAEKGTPEEATAB---AYRYADLLAKENGKTVV 57  
 Db 296 KAQLPKKEELVILEGIVEATG--MTCQARSSY--ITSEILWGYRFTPLVLTWEDGFYEV 351  
 QY 58 D 58  
 Db 352 D 352  
 RESULT 13  
 IRK6\_RAT  
 ID IRK6\_RAT  
 AC P48550; STANDARD; PRT; 425 AA.  
 DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G protein-activated inward rectifier potassium channel 2 (GIRK2)  
 DE (Potassium channel, inwardly rectifying, subfamily J, member 6)  
 DE (Inward rectifier K+ channel Kir3.2) (KATP-2) (BIR1).  
 GN KCNJ6 OR KCNJ7 OR GIRK2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]\_TaxID=10116;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=95324735; PubMed=7601286;  
 RA Bond C.T., Aemmelae C., Ashfield R., Blair T.A., Gribble F.,  
 RA Khan R.N., Lee K., Proks P., Rowe I.C.M., Sakura H., Ashford M.J.,  
 RA Adelman J.P., Ashcroft P.M.;  
 RT "Cloning and functional expression of the cDNA encoding an inwardly-  
 RT rectifying potassium channel expressed in pancreatic beta-cells and  
 RT in the brain.";  
 RL FEBS Lett. 367:61-66(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95352112; PubMed=7626127;  
 RA Stoffel M., Tokuyama Y., Trabb J.B., German M.S., Tsaar M.L.,  
 RA Jan L.Y., Polonsky K.S., Bell G.I.;  
 RT "Cloning of rat KATP-2 channel and decreased expression in pancreatic  
 RT islets of male Zucker diabetic fatty rats.";  
 RL Biochem. Biophys. Res. Commun. 212:894-899(1995).  
 CC -!- FUNCTION: THIS POTASSIUM CHANNEL IS CONTROLLED BY G PROTEINS. IT  
 CC MAY BE INVOLVED IN THE REGULATION OF INSULIN SECRETION BY GLUCOSE  
 CC AND/OR NEUROTRANSMITTERS. INWARD RECTIFIER K+ CHANNELS ARE  
 CC CHARACTERIZED BY A GREATER TENDENCY TO ALLOW POTASSIUM TO FLOW  
 CC INTO THE CELL RATHER THAN OUT OF IT. THEIR VOLTAGE DEPENDANCE IS  
 CC REGULATED BY THE CONCENTRATION OF EXTRACELLULAR POTASSIUM; AS  
 CC EXTERNAL K+ IS RAISED, THE VOLTAGE RANGE OF THE CHANNEL OPENING  
 CC SHIFTS TO MORE POSITIVE VOLTAGES. THE INWARD RECTIFICATION IS  
 CC MAINLY DUE TO THE BLOCKAGE OF OUTWARD CURRENT BY INTERNAL  
 CC MAGNESIUM. CAN BE BLOCKED BY EXTERNAL BA2+ OR CS+.  
 CC -!- SUBUNIT: ASSOCIATES WITH GIRK1 OR GIRK4 TO FORM A G-PROTEIN-  
 CC ACTIVATED HETEROMULTIMER PORE-FORMING UNIT. THE RESULTING INWARD  
 CC CURRENT IS MUCH LARGER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: PANCREATIC BETA-CELLS AND BRAIN.  
 CC -!- SIMILARITY: BELONGS TO THE INWARD RECTIFIER-TYPE K+ CHANNEL  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X83583; CAA58566.1; --  
 CC EMBL; U21087; AAA87002.1; --  
 DR PIR; S66266; S52852.  
 DR InterPro; IPR001838; K-channel\_IR.  
 DR Pfam; PF01007; IRK; 1.  
 DR PRINTS; PR01320; KIRCHANNEL.  
 DR ProDom; PD001103; K-channel\_IR; 2.  
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;  
 KW Potassium transport.  
 FT DOMAIN 1 96 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 97 119 M1 (POTENTIAL).  
 FT DOMAIN 144 160 H5 (POT-FORMING) (POTENTIAL).  
 FT TRANSMEM 169 193 M2 (POTENTIAL).  
 FT DOMAIN 194 425 CYTOPLASMIC (POTENTIAL).  
 FT SITE 184 184 ROLE IN THE CONTROL OF POLYAMINE-MEDIATED CHANNEL GATING AND IN THE BLOCKING BY INTRACELLULAR MAGNESIUM (BY SIMILARITY).  
 FT FT







```
FT DOMAIN 6275 6368 FIBRONECTIN TYPE-III.
FT DOMAIN 6413 6502 IG-LIKE C2-TYPE 48.
FT DOMAIN 6507 6596 IG-LIKE C2-TYPE 49.
FT DISULFID 568 621 POTENTIAL.
FT DISULFID 2908 2975 POTENTIAL.
FT DISULFID 3015 3065 POTENTIAL.
FT DISULFID 3707 3759 POTENTIAL.
FT DISULFID 3826 3890 POTENTIAL.
FT DISULFID 5092 5157 POTENTIAL.
FT DISULFID 5298 5350 POTENTIAL.
FT DISULFID 5508 5560 POTENTIAL.
FT DISULFID 5616 5669 POTENTIAL.
FT DISULFID 5722 5764 POTENTIAL.
FT DISULFID 5836 5901 POTENTIAL.
FT DISULFID 5946 5998 POTENTIAL.
FT DISULFID 6036 6171 POTENTIAL.
FT DISULFID 6421 6486 POTENTIAL.
FT CONFLICT 2137 2137 A -> P (IN REF. 1).
FT CONFLICT 2245 2247 AKA -> PKP (IN REF. 1).
FT CONFLICT 2258 2258 A -> P (IN REF. 1).
FT CONFLICT 2284 2284 E -> G (IN REF. 1).
FT CONFLICT 2297 2297 M -> I (IN REF. 1).
FT CONFLICT 3531 3531 A -> G (IN REF. 1).
FT CONFLICT 3884 3888 DAGEY -> RRRRI (IN REF. 1).
FT CONFLICT 3929 3929 A -> V (IN REF. 1).
FT CONFLICT 5134 5134 A -> P (IN REF. 1).
FT CONFLICT 5145 5145 T -> S (IN REF. 1).
FT CONFLICT 5185 5185 G -> A (IN REF. 1).
FT CONFLICT 5199 5199 K -> N (IN REF. 1).
FT CONFLICT 5202 5202 L -> F (IN REF. 1).
FT CONFLICT 5213 5213 F -> L (IN REF. 1).
FT CONFLICT 6178 6178 A -> G (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SQ SEQUENCE 6632 AA; 731665 MW; 262D3EDD62960E89 CRC64;
```

```
Query Match 17.1%; Score 63; DB 1; Length 6632;
Best Local Similarity 30.6%; Pred. No. 4.9e+02;
Matches 26; Conservative 12; Mismatches 33; Indels 14; Gaps 4;

Qy 2 EKTPEEPKEVTKANLIY----ADGK-TQTAEFKGTFEATAEAYRYADLLAKENGK-- 54
Db 2112 EKIVESRTITTIKTEDVYTLKISNAKIEQTGTVKVTQNSAGQDSKQADLKVEPNVKAP 2171

Qy 55 -----YTVDVADKGYTL--NIKPAQ 72
Db 2172 KFKSQTDRKVADEGEPLRWNLDDG 2196
```

Search completed: September 3, 2003, 11:20:06  
Job time : 11.1233 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:14:37 ; Search time 17.0137 Seconds  
(without alignments)  
406.975 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_153\_224  
Perfect score: 369  
Sequence: 1 KEKTPPEPKKEVTIKANLIY.....GKYTVDVADKGYTLNIKFPAG 72

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	369	100.0	455	2 A45063	immunoglobulin-bin
2	369	100.0	719	2 A42808	Ig light chain-bin
3	315	85.4	992	2 S54396	protein L precursor
4	162	43.9	74	2 A34483	Ig light chain-bin
5	82.5	22.4	1576	2 AE0249	probable hemolysin
6	80.5	21.8	395	2 A43700	BNS1 protein - hum
7	74	20.1	871	2 D86355	protein T16E15.12
8	73.5	19.9	256	2 S54803	merozoite surface
9	72	19.5	1025	2 E86355	hypothetical prote
10	69	18.7	324	2 T37931	hypothetical coile
11	68	18.4	448	2 B44782	late embryogenesis
12	67	18.2	953	1 H30169	leukotoxin A - Pas
13	66.5	18.0	657	2 AD1525	probable cell surf
14	66	17.9	448	2 JC6171	late embryogenesis
15	66	17.9	470	2 AF2828	PAD dependent oxid
16	66	17.9	477	2 D97606	glycolate oxidase
17	65.5	17.8	436	2 T16638	hypothetical prote
18	65.5	17.8	476	1 W2EC	tryptophanase (EC
19	65.5	17.8	476	2 E91209	tryptophanase [imp
20	65.5	17.8	476	2 H86055	tryptophanase [imp
21	65.5	17.8	1179	2 T35093	DNA-directed DNA p
22	65.5	17.8	4558	2 C82199	RTX toxin RtxA Vcl
23	65	17.6	636	2 H70184	methyl-accepting c
24	65	17.6	838	2 D71492	hypothetical prote
25	64.5	17.5	621	2 F72719	hypothetical prote
26	64	17.3	621	2 S49020	nuclear lamin C pr
27	64	17.3	722	2 T21521	hypothetical prote
28	63.5	17.2	414	2 S48738	potassium channel
29	63.5	17.2	425	2 S52852	inward rectifier p

30	63.5	17.2	479	2 T47561	late embryogenesis
31	63	17.1	88	2 C83657	hypothetical prote
32	63	17.1	319	2 E89872	hypothetical prote
33	63	17.1	693	2 H95255	choline binding pr
34	63	17.1	6642	2 T29757	protein UNC-89 - C
35	62.5	16.9	423	2 I38979	inward rectifier p
36	62.5	16.9	425	2 I48202	potassium channel-
37	62	16.8	118	2 S38717	Ig heavy chain V r
38	62	16.8	195	2 A70247	conserved hypotet
39	62	16.8	356	2 T37136	hypothetical prote
40	62	16.8	529	2 S62194	hypothetical prote
41	62	16.8	1612	2 A81347	probable peptidogl
42	61.5	16.7	266	2 A12289	hypothetical prote
43	61.5	16.7	280	2 S35103	bone sialoprotein
44	61.5	16.7	451	1 D64424	tLd homolog MJ099
45	61.5	16.7	487	2 T10215	hypothetical prote

ALIGNMENTS

RESULT 1

A45063

immunoglobulin-binding protein LG - Peptostreptococcus magnus

C:Species: Peptostreptococcus magnus

C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999

C:Accession: A45063

R:Kihlberg, B.M.; Sjoerding, U.; Kastern, W.; Bjoerck, L.

J. Biol. Chem. 267, 25583-25588, 1992

A>Title: Protein LG: a hybrid molecule with unique immunoglobulin binding properties.

A:Reference number: A45063; MUID:93094283; PMID:1460053

A:Accession: A45063

A>Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-455 <KH>

A:Cross-references: GB:S50809; NID:G261705; PIDN:AAA03280.1; PID:G261706

A>Note: sequence extracted from NCBI backbone (NCBIN:120302, NCBIPI:120303)

Query Match 100.0%; Score 369; DB 2; Length 455;

Best Local Similarity 100.0%; Pred. No. 6.8e-31;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAFKGTPEATAEAYRYADLLAKENGKVTVDVA 60

Db 174 KEKTPPEPKKEVTIKANLIYADGKTQTAFKGTPEATAEAYRYADLLAKENGKVTVDVA 233

QY 61 DKGVTLNKFPAG 72

Db 234 DKGVTLNKFPAG 245

RESULT 2

A42808

Ig light chain-binding protein precursor - Peptostreptococcus magnus

N:Alternate names: protein L

C:Species: Peptostreptococcus magnus

C>Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 07-May-1999

C:Accession: A42808; A41493

R:Kastern, W.; Sjoerding, U.; Bjoerck, L.

J. Biol. Chem. 267, 12820-12825, 1992

A>Title: Structure of peptostreptococcal protein L and identification of a repeated immu

A:Reference number: A42808; MUID:92316971; PMID:1618782

A:Accession: A42808

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-719 <KAS>

A:Cross-references: GB:M86697

R:Kastern, W.; Holst, E.; Nielsen, E.; Sjoerding, U.; Bjoerck, L.

Infect. Immun. 58, 1217-1222, 1990

A>Title: Protein L, a bacterial immunoglobulin-binding protein and possible virulence de

A:Reference number: A41493; MUID:90215984; PMID:2108927

A:Accession: A41493

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 202-275 <KA2>

C:Keywords: immunoglobulin

Query Match 100.0%; Score 369; DB 2; Length 719;  
Best Local Similarity 100.0%; Pred. No. 1.1e-30;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDVA 60  
|||||  
Db 246 KEKTPPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDVA 305  
|||||

QY 61 DRGYTLNIKFPAG 72

Db 306 DRGYTLNIKFPAG 317

RESULT 3

S54396

protein L precursor - Peptostreptococcus magnus (strain 3316)

C:Species: Peptostreptococcus magnus

A:Variety: strain 3316

C>Date: 27-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 15-Oct-1999

C:Accession: S54396

R:Murphy, J.P.; Duggleby, C.J.; Atkinson, M.A.; Trowern, A.R.; Atkinson, T.; Goward, C.R.

Mol. Microbiol. 12, 911-920, 1994

A:Title: The functional units of a peptostreptococcal protein L.

A:Reference number: S54396; MUID:95020613; PMID:7934898

A:Accession: S54396

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-992 <MUR>

A:Cross-references: EMBL:L04466; NID:gl50673; PIDN:AAA67503.1; PID:gl50674

Query Match

Best Local Similarity 85.4%; Score 315; DB 2; Length 992;

Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTEPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDVADK 62  
|||||  
Db 468 ETEPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDLEDG 527  
|||||

QY 63 GYTINIKFPAG 72

Db 528 GYTINIKFPAG 537

RESULT 4

A34483

Ig light chain-binding protein L - Peptostreptococcus magnus (fragments)

C:Species: Peptostreptococcus magnus

C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 16-Aug-1996

C:Accession: A34483

R:Akerstroem, B.; Bjoerck, L.

J. Biol. Chem. 264, 19740-19746, 1989

A:Title: Protein L: an immunoglobulin light chain-binding bacterial protein. Characteriz

A:Reference number: A34483; MUID:90062074; PMID:2479638

A:Accession: A34483

A:Molecule type: protein

A:Residues: 1-74 <AKB>

C:Keywords: immunoglobulin

Query Match 43.9%; Score 162; DB 2; Length 74;

Best Local Similarity 62.5%; Pred. No. 4e-10; 2; Indels 16; Gaps 1;

Matches 35; Conservative 3; Mismatches 2; Indels 16; Gaps 1;

QY 2 EKTPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDVA 41  
|||||

Db 9 ETTPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDVA 64  
|||||

RESULT 5

AS0249

protein T16E15.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: D86355

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

probable hemolysin YPO2045 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AS0249

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AS0001; MUID:21470413; PMID:11586360

A:Accession: AS0249

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1576 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90857.1; PID:gl5980056; GSPDB:GN00175

C:Genetics:

A:Gene: YPO2045

Query Match

Best Local Similarity 22.4%; Score 82.5; DB 2; Length 1576;

Matches 30; Conservative 9; Mismatches 27; Indels 43; Gaps 4;

QY 1 KEKTPPEKPEVTIKANLIYADGKQTQAEFGKTPPEATAEAYRYADLLAKENGKTYTDVA 36  
|||||

Db 1001 KANTTEQEKGEVSLRG-----GMTATQEIKGHLGVKAETSGDSYAEMLVGNINAKSG 1054  
|||||

QY 37 ----ATAEAYRYADLLAKENGKTYTD-----VADKGYTLNIKFPAG 72  
|||||

Db 1055 VSIKTTGDYYYATNIEGGNGDITIDAGNNLYFDQVQDSQRSSNIKFSG 1103  
|||||

RESULT 6

A43700

EN51 protein - human

C:Species: Homo sapiens (man)

C>Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 21-Jul-2000

C:Accession: A43700

R:Ittmann, M.; Greco, A.; Basilico, C.

Mol. Cell. Biol. 7, 3386-3393, 1987

A:Title: Isolation of the human gene that complements a temperature-sensitive cell cycle

A:Reference number: A43700; MUID:88065472; PMID:3683386

A:Accession: A43700

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-395 <ITT>

A:Cross-references: GB:ML7754; NID:gl79512; PIDN:AAA51838.1; PID:gl79513

C:Genetics:

A:Gene: GDB:BN51T

A:Cross-references: GDB:119728; OMIM:187280

A:Map position: 8pter-8q24

Query Match

Best Local Similarity 21.8%; Score 80.5; DB 2; Length 395;

Matches 26; Conservative 9; Mismatches 27; Indels 19; Gaps 3;

QY 3 KTEPEKPEVTIKANLIYAD-----GKTQTAEFGKTPPEATAEAYRYADLLAK 50  
|||||

Db 63 KKEPEKPEVTIKKEKREDRDRQREGHGRRRRPEVIOQSHSIFEQPAEMWK-----KK 117  
|||||

QY 51 ENGYTYDVAADKG--YTANIK 69  
|||||

Db 118 GNWDKTVDSVDMGSPSHINIK 138  
|||||

RESULT 7

D86355

protein T16E15.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: D86355

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000  
C.A.; Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; PMID:21016719; PMID:11130712

A;Cross-references: GB:AE005172; NID:g9392688; PIDN:AAF87265.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: Tf6E15.12  
A;Map position: 1

Query Match 20.1%; Score 74; DB 2; Length 871;  
Best Local Similarity 26.7%; Pred. No. 9.6;  
Matches 20; Conservative 16; Mismatches 27; Indels 12; Gaps 2;

Qy 3 KTPPEPKREVITI-----KANLIYADGKTQTAFKGTPEATAEAYRVADLLAKE 51  
|||:::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db 429 KTSEDKKQELSLKLSSLEMSKECKEKLDADRQVEELE-TLQKSESHQLDALLAKE 487

Qy 52 NGKYTVDVVDAGGYTL 66  
: : : : :  
Db 488 VNQLQTVIEEGHVI 502

RESULT 8  
S54803  
merozoite surface glycoprotein precursor - Theileria parva (fragment)  
C;Species: Theileria parva  
C;Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Sep-1997  
C;Accession: S54803  
R/Shields, B.R.; d'Oliviera, C.; McKellar, S.; Ben-Miled, L.; Kawazu, S.; Hide, G.  
submitted to the EMBL Data Library, March 1995  
A;Description: Selection of diversity at putative glycosylation sites in the immunodominant  
A;Reference number: S54803  
A;Accession: S54803  
A>Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-256 <SHI>  
A;Cross-references: EMBL:S54803; NID:g807986; PID:g807987  
C;Keywords: glycoprotein

Query Match 19.9%; Score 73.5; DB 2; Length 256;  
Best Local Similarity 31.7%; Pred. No. 2.8;  
Matches 19; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

Qy 1 KEKTPPEPKREVITIKANLIYADGKTQTAFKGTPEATAEAYRVADLLAKENKGYTVDVA 60  
:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 25 EEKKKKKKEDLVIVDTLSSWENVTSPEAGGTLLKAN-EGYRPFTLKVGDKTYLNVDTS 83

RESULT 9  
E86355  
hypothetical protein Tf6E15.11 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C;Accession: E86355  
R/Theologis, A.; Eckler, J.R.; Palm, C.J.; Federpietel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
C.A.; Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; PMID:21016719; PMID:11130712

probable cell surface protein (LPXTG motif) - Listeria innocua (strain Clip1)

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C;Accession: ADI525

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
; Dominguez-Bernal, G.; Duhaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunat, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A.;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; UID:21537279; PMID:11679669

A;Accession: ADI525

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-657 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC95972.1; PID:g16413192; GSPDB:GN00178

A;Experimental source: strain Clip1262

C;Genetics:

A;Gene: lin0740

Query Match 18.0%; Score 66.5; DB 2; Length 657;  
Best Local Similarity 32.5%; Pred.No.43;  
Matches 26; Conservative 6; Mismatches 29; Indels 19; Gaps 5;

QY 1 KEKTPPEPKEEVTKANLIYADGK----TQTAEFKG-----TFEATAEAAYRVADLLAK 50  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
381 KDPVPAKP---VTVK---YVDKGKELAPSETFTGFIDDDYTSTETKTIEGYTLVETPAN 433

Db

QY 51 ENKGYTVD--VADKGYTLNI 68  
||| ||| : ||| |||  
434 ANGKLTAQQQTVNYIYTKNI 453

Db

RESULT 14

JC6171

late embryogenesis abundant protein - Arabidopsis thaliana

N;Alternate names: embryogenic cell protein 63; phosphotyrosine protein

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 03-Jun-2002

C;Accession: JC6171

R;Yang, H.; Saitou, T.; Komeda, Y.; Harada, H.; Kamada, H.

Gene 184, 83-88, 1997

A;Title: Arabidopsis thaliana ECP63 encoding a LEA protein is located in chromosome 4.

A;Reference number: JC6171; UID:97169149; PMID:9016956

A;Accession: JC6171

A;Molecule type: mRNA

A;Residues: 1-448 <XAN>

A;Cross-references: DDBJ:D64140; NID:g1526423; PIDN:BAAI1017.1; PID:g1526424

C;Comment: This protein is a phosphotyrosine protein which belongs to late embryogenesis

C;Genetics:

A;Gene: ATECP63

A;Map position: 4, south part

C;Superfamily: pea seed biotin-containing protein

C;Keywords: seed

Query Match 17.9%; Score 66; DB 2; Length 448;  
Best Local Similarity 26.4%; Pred.No.32;  
Matches 23; Conservative 16; Mismatches 32; Indels 16; Gaps 3;

QY 1 KEKTPPEPKEEVTKANLIYADGKTQTATFEKGTPEATAE-----AYRYADLL 48  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
176 KDTAEKAKESNTYADKA-KKAKDKTAEKVGSEKYDYTVDKAVEARDYTAEKIAEKDKT 234

Db

QY 49 AKENGK---YTVDVADKGYTLNIKFAAG 72  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
235 AEKTYGEYKDYTVKEATEGKDVTVSKLG 261

Db

RESULT 15

AF2828

FAD dependent oxidoreductase Atu2053 [imported] - Agrobacterium tumefaciens (strain C58,

C;Species: Agrobacterium tumefaciens

Search completed: September 3, 2003, 11:24:31  
Job time : 19.0137 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: September 3, 2003, 11:14:02 ; Search time 44.137 Seconds  
(without alignments)  
420.957 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_153\_224

Perfect score: 369

Sequence: 1 KEKTPPEPKEEVTIKANLIY.....GKYTVADVADGKGYTLNIKFKAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriaph:\*

17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	369	100.0	455	2 Q53291	Q53291 streptococ
2	369	100.0	719	2 Q51912	Q51912 peptostrept
3	315	85.4	992	2 Q51918	Q51918 peptostrept
4	88.5	24.0	216	4 Q3BPV7	Q3BPV7 homo sapien
5	88.5	24.0	398	4 Q9BPZ1	Q9BPZ1 homo sapien
6	83.5	22.6	398	11 Q9CZ02	Q9CZ02 mus musculus
7	83.5	22.6	398	11 Q91WD1	Q91WD1 mus musculus
8	82.5	22.4	1576	16 Q8ZEV8	Q8ZEV8 yersinia pe
9	74	20.1	871	10 Q9LME2	Q9LME2 arabisopsis
10	73.5	19.9	256	5 Q27039	Q27039 theileria p
11	73.5	19.9	256	5 Q27029	Q27029 theileria p
12	73.5	19.9	280	5 Q27030	Q27030 theileria p
13	72	19.5	1025	10 Q9LME3	Q9LME3 arabisopsis
14	71	19.2	265	10 Q23822	Q23822 dunaliella
15	69	18.7	324	3 Q5UUK3	Q5UUK3 schizosacch
16	68	18.4	448	10 Q9SKP0	Q9SKP0 arabisopsis

17	67.5	18.3	383	5 Q9VTV4	Q9vtv4 drosophila
18	67	18.2	946	2 Q9EV24	Q9ev24 mannheimia
19	67	18.2	953	2 Q9ETX2	Q9etx2 mannheimia
20	67	18.2	953	2 Q9EV29	Q9ev29 pasteurella
21	67	18.2	953	2 Q9EV25	Q9ev25 mannheimia
22	67	18.2	953	2 Q9ETG5	Q9etg5 pasteurella
23	67	18.2	953	2 Q9EV23	Q9ev23 mannheimia
24	67	18.2	953	2 Q9EV34	Q9ev34 pasteurella
25	67	18.2	1204	2 Q8GM76	Q8gm76 haemophilus
26	66.5	18.0	657	16 Q92DS2	Q92ds2 listeria in
27	66	17.9	448	10 Q96246	Q96246 arabidopsis
28	66	17.9	477	16 Q8UDR8	Q8udr8 agrobacteri
29	66	17.9	1096	2 Q8GM79	Q8gm79 haemophilus
30	66	17.9	1098	2 Q48152	Q48152 haemophilus
31	66	17.9	1498	5 Q8IDP2	Q8idp2 plasmodium
32	65.5	17.8	436	5 Q21481	Q21481 caenorhabdi
33	65.5	17.8	451	2 Q924J9	Q924j9 lactobactil
34	65.5	17.8	476	16 Q8FBV2	Q8fbv2 escherichia
35	65.5	17.8	1167	17 Q8TUJ9	Q8tuj9 methanosarc
36	65.5	17.8	1210	2 Q8GM75	Q8gm75 haemophilus
37	65.5	17.8	1210	2 Q8GM74	Q8gm74 haemophilus
38	65.5	17.8	4545	2 Q9X4W2	Q9x4w2 vibrio chol
39	65.5	17.8	4558	16 Q9KS12	Q9ks12 vibrio chol
40	65	17.6	284	5 Q76174	Q76174 theileria s
41	65	17.6	284	5 Q9TYA5	Q9tya5 theileria s
42	65	17.6	636	16 Q51624	Q51624 borrelia bu
43	65	17.6	838	16 Q84625	Q84625 chlamydia t
44	64.5	17.5	243	17 Q9YFE6	Q9yfe6 aeropyrum p
45	64.5	17.5	333	3 Q9UVE8	Q9uve8 yarrowia li

#### ALIGNMENTS

#### RESULT 1

ID	Q53291	PRELIMINARY;	PRT;	455 AA.
AC	Q53291;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Protein LG (Fragment).			
OS	Streptococcus sp.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1306;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=93094283; PubMed=1460053;			
RA	Kihlberg B.M., Sjobring U., Kastern W., Bjorck L.;			
RT	"Protein LG: a hybrid molecule with unique immunoglobulin binding			
RT	properties.";			
RL	J. Biol. Chem. 267:25583-25588(1992).			
DR	EMBL; S50809; AAA03280.1; -			
DR	HSSP; P06654; IFGX.			
DR	InterPro; IPR003147; BL.			
DR	InterPro; IPR000724; Igg_bind_B.			
DR	Pfam; PF02246; BL; 4.			
DR	Pfam; PF01378; Igg_binding_B; 2.			
FT	NON_TER 455 455			
SQ	SEQUENCE 455 AA; 49926 MW; 381FC235B8C307B CRC64;			

Query Match 100.0%; Score 369; DB 2; Length 455;

Best Local Similarity 100.0%; Pred. No. 2.4e-29;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KEKTPPEPKEEVTIKANLIYADGKGYTLNIKFKAGTQTAETAEFGTTEATAEAYRAYADLLAKENGKGYTVDA 60

Db 174 KEKTPPEPKEEVTIKANLIYADGKGYTLNIKFKAGTQTAETAEFGTTEATAEAYRAYADLLAKENGKGYTVDA 233

Qy 61 DKGYTLNIKFKAG 72

Db 234 DKGYTLNIKFKAG 245

```

RESULT 2
ID Q51912 PRELIMINARY; PRT; 719 AA.
AC Q51912;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=90215984; PubMed=2108927;
RA Kastern W., Holst E., Nielsen E., Sjoerding U., Bjorck L.;
RT "Protein L, a bacterial immunoglobulin-binding protein and possible
RT virulence determinant.";
RL Infect. Immun. 58:1217-1222(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=92316971; PubMed=1618782;
RA Bjoerck L., Sjoerding U., Kastern W.;
RT "Structure of peptostreptococcal protein L and identification of
RT repeated immunoglobulin light chain-binding domain.";
RL J. Biol. Chem. 267:12820-12825(1992).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; M86697; AAA25612.1; -
DR InterPro: IPR003147; B1.
DR InterPro: IPR004829; Csurface_antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 5.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR ProDom: PD153432; Csurface_antigen; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 719 PROTEIN L.
SQ SEQUENCE 719 AA; 78983 MW; 963A8D76D5E34DD2 CRC64;

Query Match 100.0%; Score 369; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 4.1e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVA 60
Db 246 KEKTPPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVA 305
QY 61 DKGYTLNLIKFG 72
Db 306 DKGYTLNLIKFG 317

RESULT 3
Q51918 PRELIMINARY; PRT; 992 AA.
ID Q51918;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=95078460; PubMed=7987012;
RA Murphy J.P., Trowern A.R., Duggleby C.J.;
RT "Nucleotide sequence of the gene for peptostreptococcal protein L.";
RL DNA Seq. 4:259-263(1994).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; L04456; AAA67503.1; -
DR HSP; Q51911; 1GAB.
DR InterPro: IPR003147; B1.
DR InterPro: IPR002988; GA.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 4.
DR Pfam: PF01468; GA; 4.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 992 PROTEIN L.
SQ SEQUENCE 992 AA; 108700 MW; 9CFF5771578A5DCE CRC64;

Query Match 85.4%; Score 315; DB 2; Length 992;
Best Local Similarity 85.7%; Pred. No. 1.9e-23;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTEPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 62
Db 468 ETEPEPEEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 527
QY 63 GYTLNLIKFG 72
Db 528 GYTLNLIKFG 537

RESULT 4
Q9BPV7 PRELIMINARY; PRT; 216 AA.
ID Q9BPV7;
AC Q9BPV7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to BNS1 (BHK21) temperature sensitivity complementing.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003039; AAH03039.1; -
DR EMBL; BC000516; AAH00516.1; -
SQ SEQUENCE 216 AA; 24806 MW; F29028AEAC8DB04 CRC64;

Query Match 24.0%; Score 88.5; DB 4; Length 216;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 27; Conservative 9; Mismatches 26; Indels 19; Gaps 3;

QY 3 KTEPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 50
Db 66 KIKEPEPEEVTIKKRRDRDRQREGHGRGRPEVIQSHSIFEQGPAEMMK-----KK 120
QY 51 ENKGYTVDVADKG--VTLNIK 69
Db 121 GNWDKTYVDVSDMGPSHIINIK 141

RESULT 5
Q9BP21 PRELIMINARY; PRT; 141 AA.
ID Q9BP21;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=95078460; PubMed=7987012;
RA Murphy J.P., Trowern A.R., Duggleby C.J.;
RT "Nucleotide sequence of the gene for peptostreptococcal protein L.";
RL DNA Seq. 4:259-263(1994).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; L04456; AAA67503.1; -
DR HSP; Q51911; 1GAB.
DR InterPro: IPR003147; B1.
DR InterPro: IPR002988; GA.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 4.
DR Pfam: PF01468; GA; 4.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 992 PROTEIN L.
SQ SEQUENCE 992 AA; 108700 MW; 9CFF5771578A5DCE CRC64;

Query Match 85.4%; Score 315; DB 2; Length 992;
Best Local Similarity 85.7%; Pred. No. 1.9e-23;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTEPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 62
Db 468 ETEPEPEEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 527
QY 63 GYTLNLIKFG 72
Db 528 GYTLNLIKFG 537

RESULT 4
Q9BPV7 PRELIMINARY; PRT; 216 AA.
ID Q9BPV7;
AC Q9BPV7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to BNS1 (BHK21) temperature sensitivity complementing.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003039; AAH03039.1; -
DR EMBL; BC000516; AAH00516.1; -
SQ SEQUENCE 216 AA; 24806 MW; F29028AEAC8DB04 CRC64;

Query Match 24.0%; Score 88.5; DB 4; Length 216;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 27; Conservative 9; Mismatches 26; Indels 19; Gaps 3;

QY 3 KTEPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 50
Db 66 KIKEPEPEEVTIKKRRDRDRQREGHGRGRPEVIQSHSIFEQGPAEMMK-----KK 120
QY 51 ENKGYTVDVADKG--VTLNIK 69
Db 121 GNWDKTYVDVSDMGPSHIINIK 141

RESULT 5
Q9BP21 PRELIMINARY; PRT; 141 AA.
ID Q9BP21;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=95078460; PubMed=7987012;
RA Murphy J.P., Trowern A.R., Duggleby C.J.;
RT "Nucleotide sequence of the gene for peptostreptococcal protein L.";
RL DNA Seq. 4:259-263(1994).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; L04456; AAA67503.1; -
DR HSP; Q51911; 1GAB.
DR InterPro: IPR003147; B1.
DR InterPro: IPR002988; GA.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 4.
DR Pfam: PF01468; GA; 4.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 992 PROTEIN L.
SQ SEQUENCE 992 AA; 108700 MW; 9CFF5771578A5DCE CRC64;

Query Match 85.4%; Score 315; DB 2; Length 992;
Best Local Similarity 85.7%; Pred. No. 1.9e-23;
Matches 60; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KTEPEPEEVTIKANLIYADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 62
Db 468 ETEPEPEEVTIKVNLIFADGKTQTAEFGKTFEATAEAYRYADLLAKENGKTYVDVADK 527
QY 63 GYTLNLIKFG 72
Db 528 GYTLNLIKFG 537

```



```

Query Match          22.6%; Score 83.5; DB 11; Length 398;
Best Local Similarity 32.1%; Pred. No. 2.2;
Matches 26; Conservative 10; Mismatches 26; Indels 19; Gaps 3;

QY 3 KTEEPKEEYTIKANLIYAD-----GKTQTAEPKGTPEEATAEAYRYADLLAK 50
   | ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 KKEEPKEEYTKMEKKERDRDRQREGHGRGRPEVIQSHSIFEQGPAAEMMK-----KK 120
   | ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 51 ENGKYTVDVADKG--YTLNIK 69
   | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GNWDKTVDSMDGSPSHIINIK 141
   | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
Q91WD1 PRELIMINARY; PRT; 398 AA.
AC Q91WD1; 2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE RIKEN CDNA 2810426M17 gene.
GN 2810426M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RL Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016102; AAH16102.1; -.
DR MGD; MGI:1914315; 2810426M17RIK.
SQ SEQUENCE 398 AA; 44323 MW; D509DE632C656FFD CRC64;

Query Match          22.6%; Score 83.5; DB 11; Length 398;
Best Local Similarity 32.1%; Pred. No. 2.2;
Matches 26; Conservative 10; Mismatches 26; Indels 19; Gaps 3;

QY 3 KTEEPKEEYTIKANLIYAD-----GKTQTAEPKGTPEEATAEAYRYADLLAK 50
   | ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 KKEEPKEEYTKMEKKERDRDRQREGHGRGRPEVIQSHSIFEQGPAAEMMK-----KK 120
   | ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 51 ENGKYTVDVADKG--YTLNIK 69
   | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GNWDKTVDSMDGSPSHIINIK 141
   | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
Q8ZEV8 PRELIMINARY; PRT; 1576 AA.
AC Q8ZEV8;
AC Q8ZEV8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative hemolysin.
GN YPO2045 OR Y2267.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
NCBI_TaxID=632;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA "Genome sequence of Yersinia pestis, the causative agent of plague."
RT

```

```

RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Millies M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL: AJ414151; CAC90857.1; -
DR EMBL: AB013829; AAM83827.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1576 AA; 169976 MW; 7029BC4BA386424C CRC64;

Query Match 22.4%; Score 82.5; DB 16; Length 1576;
Best Local Similarity 27.5%; Pred. No. 14;
Matches 30; Conservative 9; Mismatches 27; Indels 43; Gaps 4;

QY 1 KEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTGTF-----EE----- 36
Db 1001 KANTPEQKEVSLRG-----GMTATQEIKGHLGVKAETSQGDSVAEEMLVGNINAKSG 1054
QY 37 ----ATAEAYRYADLLAKENGKYTVD-----VADKGYTLNKFAG 72
Db 1055 VSIKTTGDAYYATNIEGNGDITDAGNNLYFDQVQDSRSSNIKFSG 1103

RESULT 9
Q9LME2 PRELIMINARY; PRT; 871 AA.
ID Q9LME2
AC Q9LME2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE T16E15-12.
DE T16E15-12.
GN T16E15.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Liu S.X., Sakano H., Etgu P., Lee J., Lenz C., Pham P., Toriumi M.,
RA Yu G., Chan A., Chung M., Goldsmith A., Liu A., Smith A., Vaysberg M.,
RA Altafi H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,
RA Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,
RA Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
RA Federspiel N.A., Theologis A.;
RT "The sequence of BAC T16E15 from Arabidopsis thaliana chromosome 1.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC068562; AAF87265.1; -
SQ SEQUENCE 871 AA; 100006 MW; E8D2778FF98CBE54 CRC64;

Query Match 20.1%; Score 74; DB 10; Length 871;
Best Local Similarity 26.7%; Pred. No. 52;
Matches 20; Conservative 16; Mismatches 27; Indels 12; Gaps 2;

QY 3 KYPEPEKEEVTI-----KANLIYADGKTQTAEFKGTGFEATAEAYRYADLLAKE 51
Db 429 KTSDEKKKESLKSLEMSKECKEKLQADQROVEELE-TLQKESHQLOADLLAKE 487
QY 52 NGKYTVDVADKGYTL 66
Db 488 VNQLQTVIEKGHVI 502

RESULT 10
Q27039 PRELIMINARY; PRT; 256 AA.
ID Q27039;
AC Q27039;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE T16E15-12.
DE T16E15-12.
GN T16E15.12.
OS Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Muguga;
RX MEDLINE=96089836; PubMed=8538686;
RA Shields B.R., d'Oliveira C., McKellar S., Ben-Miled L., Kawazu S.,
RA Hide G.;
RT "Selection of diversity at putative glycosylation sites in the
RT immunodominant merozoite/piroplasm surface antigen of Theileria
RT parasites.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z48740; CAA88632.1; -
DR InterPro: IPR003407; Merozoite_Agen.
DR Pfam: PF02488; EMA; 1.
KW Signal.
FT SIGNAL 1 23
FT NON_TER 256 256
SQ SEQUENCE 256 AA; 29528 MW; C982AE71A45DC5B CRC64;

Query Match 19.9%; Score 73.5; DB 5; Length 256;
Best Local Similarity 31.7%; Pred. No. 14;
Matches 19; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

QY 1 KEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTGFEATAEAYRYADLLAKENGKYTVDVA 60
Db 25 EEEKKKEDLTVDVTLSSWENVTSTPEAGGTLLKAN-EGYRFKTLKVGDKTLYNDTS 83

RESULT 11
Q27029 PRELIMINARY; PRT; 256 AA.
ID Q27029;
AC Q27029;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE T16E15-12.
DE T16E15-12.
GN T16E15.12.
OS Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Muguga;
RX MEDLINE=96089836; PubMed=8538686;
RA Shields B.R., d'Oliveira C., McKellar S., Ben-Miled L., Kawazu S.,
RA Hide G.;
RT "Selection of diversity at putative glycosylation sites in the
RT immunodominant merozoite/piroplasm surface antigen of Theileria
RT parasites.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z48740; CAA88632.1; -
DR InterPro: IPR003407; Merozoite_Agen.
DR Pfam: PF02488; EMA; 1.
KW Signal.
FT SIGNAL 1 23
FT NON_TER 256 256
SQ SEQUENCE 256 AA; 29528 MW; C982AE71A45DC5B CRC64;

Query Match 19.9%; Score 73.5; DB 5; Length 256;
Best Local Similarity 31.7%; Pred. No. 14;
Matches 19; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

QY 1 KEKTPPEKPEEVTIKANLIYADGKTQTAEFKGTGFEATAEAYRYADLLAKENGKYTVDVA 60

```

THE  
FOLLOWING  
TABLE  
GIVES  
THE  
NUMBER  
OF  
CARS  
AND  
TRUCKS  
REGISTERED  
IN  
THE  
STATE  
OF  
NEW  
JERSEY  
FOR  
THE  
YEARS  
1960  
THROUGH  
1964.

```

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA McDougall R.C., Rajandream M.A., Barrell B.G., Bothe G., Pohl T.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL109820; CAB52567.1; -.
DR GeneDB.SPombe: SPAC1952.03; -.
DR InterPro: IPR003323; OTU.
DR Pfam: PF02338; OTU; 1.
DR PROSITE: PS50802; OTU; 1.
SQ SEQUENCE 324 AA; 37516 MW; 48BBB435960D4B9C CRC64;

Query Match      18.7%; Score 69; DB 3; Length 324;
Best Local Similarity 28.6%; Pred. No. 52;
Matches 18; Conservative 15; Mismatches 26; Indels 4; Gaps 1;

QY 1 KKTPEPKPEVTIKANLIYADGKTQTAEFGTFTFEATAEAYRYADLLAKENGKVTVDVA 60
Db 126 KENTPOPKKSRNRQKERL---ERRKAEMKKMQAELESEKMDLKNEKKKFSKILE 181

QY 61 DKG 63
Db 182 EAG 184

Search completed: September 3, 2003, 11:23:13
Job time : 45.137 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:14:37 ; Search time 17.0137 Seconds  
(without alignments)  
406.975 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_225\_296

Perfect score: 370

Sequence: 1 KEKTPPEPKKEVTIKANLIY.....GKYTADLEDGGYTINIRFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	370	100.0	455	2 A45063	immunoglobulin-bin
2	370	100.0	719	2 A42808	Ig light chain-bin
3	329	88.9	992	2 S54396	protein L precursor
4	157	42.4	74	2 A34183	Ig light chain-bin
5	80.5	21.8	1576	2 AE0249	probable hemolysin
6	74	20.0	324	2 T37931	hypothetical colle
7	71.5	19.3	657	2 AD1525	probable cell surf
8	71	19.2	871	2 D86355	protein T16E15.12
9	70.5	19.1	256	2 S54803	merozoite surface
10	69.5	18.8	2082	2 G96602	probable receptor
11	69	18.6	337	2 AC1498	specificity determ
12	69	18.6	401	2 AE1498	specificity determ
13	69	18.6	1025	2 E86355	hypothetical prote
14	68.5	18.5	858	2 E96602	hypothetical prote
15	67	18.1	283	2 S69639	hypothetical prote
16	66.5	18.0	1029	2 F96602	hypothetical prote
17	66	17.8	495	2 D89808	hypothetical prote
18	65.5	17.7	4558	2 C82199	RTX toxin RtxA VCL
19	64.5	17.4	1179	2 T35093	DNA-directed DNA p
20	64	17.3	863	2 S06017	neuraxin - rat
21	64	17.3	2364	2 A56577	microtubule-associ
22	64	17.3	2484	1 QRMSP1	microtubule-associ
23	63.5	17.2	395	2 A43700	BN51 protein - hum
24	63.5	17.2	550	2 F75186	thermosome, chain
25	63	17.0	356	2 T37136	hypothetical prote
26	63	17.0	490	2 F38462	S-mephenytoin 4'-h
27	63	17.0	1104	1 A36866	microbial collagen
28	62.5	16.9	215	2 AC1156	transcription regu
29	62	16.8	266	2 A12289	hypothetical prote

#### RESULT 1

A45063

immunoglobulin-binding protein LG - Peptostreptococcus magnus

C;Species: Peptostreptococcus magnus

C;Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999

C;Accession: A45063

R;Kihlberg, B.M.; Sjoerding, U.; Kastern, W.; Bjorck, L.

J. Biol. Chem. 267, 25583-25588, 1992

A;Title: Protein LG: a hybrid molecule with unique immunoglobulin binding properties.

A;Reference number: A45063; MUID:93094283; PMID:1460053

A;Accession: A45063

A;Status: preliminary

A;Molecule type: mRNA; protein

A;Residues: 1-455 <KIH>

A;Cross-references: GB:S50809; NID:g261705; PID:AAA03280.1; PID:g261706

A;Note: sequence extracted from NCBI backbone (NCBIN:120302, NCBIPI:120303)

Query Match 100.0%; Score 370; DB 2; Length 455;

Best Local Similarity 100.0%; Pred. No. 4.3e-32;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFAEATAYADLLAKENGYTADLE 60

|||||

Db 246 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFAEATAYADLLAKENGYTADLE 305

|||||

QY 61 DGGYTINIRFAG 72

|||||

Db 306 DGGYTINIRFAG 317

#### RESULT 2

A42808

Ig light chain-binding protein precursor - Peptostreptococcus magnus

N;Alternate names: protein L

C;Species: Peptostreptococcus magnus

C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 07-May-1999

C;Accession: A42808; A41493

R;Kastern, W.; Sjoerding, U.; Bjorck, L.

J. Biol. Chem. 267, 12820-12825, 1992

A;Title: Structure of peptostreptococcal protein L and identification of a repeated i

A;Reference number: A42808; MUID:92316971; PMID:1618782

A;Accession: A42808

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-719 <KAS>

A;Cross-references: GB:M86697

R;Kastern, W.; Holst, E.; Nielsen, E.; Sjoerding, U.; Bjorck, L.

Infect. Immun. 58, 1217-1222, 1990

A;Title: Protein L, a bacterial immunoglobulin-binding protein and possible virulence

A;Reference number: A41493; MUID:90215984; PMID:2108927

A;Accession: A41493

A;Status: preliminary

[illegible]

Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak-  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A.:Title: Comparative genomics of *Listeria species*.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1525  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-657 <GIA>  
A:Cross-references: GB:AL502022; PIDN:CAC95972.1; PID:g16413192; GSPDB:GN00178  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin0740

```
Query Match      19.3%; Score 71.5; DB 2; Length 657;
Best Local Similarity 33.8%; Pred. No. 9.1;
Matches 27; Conservative 6; Mismatches 28; Indels 19; Gaps 5;

Qy   1 KETPEEPKEEVITKANLIYADGK-----TQTAEFG-----TFAEATAEAYRYADILLAK 50
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   381 KDVPKAPK---VTVVK----YVDKGKELAPSETFTGFIDDDYTSTEXTIEGYTLVETPAN 433

Qy   51 ENKPYATADLEDGG--YTINI 68
      ||| ||| |
Db   434 ANGKLTAADOOTVNVIYTKNI 453
```

RESULT 8  
D86355  
protein Tt6E15.12 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D86355  
Chin, R.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141. MUID:21016719. PMID:11130712

A;Residues: 1-871 <STO>  
A;Cross-references: GB:AE005172; NID:g9392688; PIDN:AAF87265.1; GSPDB:GNO0141  
C;Genetics:  
A;Gene: T16E15.12  
A;Map position: 1

Query Match	19.2%	Score 71;	DB 2;	Length 871;
Best Local Similarity	28.0%;	Pred. No. 14;		
Matches 21;	Conservative	14;	Mismatches	28;
Indels	12;	Gaps	2;	

```
Qy 3 K T P E P K E E V T I -----K A N L I Y A D G K T Q T A E F K G T F A E A T A E A Y R A D L L A K E 51
    || | | | : | : |
Db 429 K T S E D K K O E L S K L S S L E M S E K E C K E K L Q A D A Q R O V E E L - T L O K E S E S H O L A D L L A K E 487
```

QY 52 NGKYTADLEDGGYTI 66  
: | : | : |  
Db 488 VNQLQTVIEEKGHVI 502

RESULT 9  
S54803  
merozoite surface glycoprotein precursor - Theileria parva (fragment)  
C:Species: Theileria parva  
C:date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Sep-1997  
C:Accession: S54803  
R:Shiels, B.R.; d'Oliveira, C.; McKellar, S.; Ben-Miled, L.; Kawazu, S.; Hide, G.

submitted to the EMBL Data Library, March 1995  
A:Description: Selection of diversity at putative glycosylation sites in the immunodominant epitope of the HIV-1 gp120 protein  
A:Reference number: S54803  
A:Accession: S54803  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-256 <SH1>  
A:Cross-references: EMBL:Z48740; NID:g807986; PID:g807987  
C:Keywords: glycoprotein

```

Query Match      19.18; Score 70.5; DB 2; Length 256;
Best Local Similarity 31.08; Pred.No. 4;
Matches 18; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

QY 1 KEKTPPKPEVITKANLIIVADGKTQFAETAEAYRYADLLAKENGYTAD 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 EEKKKEKKEDLVDTLSSEWNTVSPKAGTTILKAN-EGYRPTKLVGDKTILYND 81

```

RESULT 10  
G96602  
probable receptor protein kinase FLG9.24 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: G96602  
R/Theologits, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.  
Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; D  
ansen, N.F.; Hughes, B.; Huizler, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marziani, A.; Rizzo, M.H.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; PMID:21016719; PMID:11130712

A:Accession: G96602  
A:Status: preliminary  
A:Molecule\_type: DNA  
A:Residues: 1-2062 <STO>  
A:Cross-references: GB:AE005173; NID:g11094719; PIDN:AAG29654.1; GSPDB:GNO0141  
C:Genetics:  
A:Gene: FlAC9.24  
A:Map position: 1

Query Match	18.8%	Score 69.5;	DB 2;	Length 2062;
Best Local Similarity	25.6%;	Pred. No. 54;		
Matches 21: Conservative	15;	Mismatches	29;	Indels 17;
Matches 21: Conservative	15;	Mismatches	29;	Indels 17;

QY	7	EPKEEVIK	ANLIYADK	TQTAEK	GTAET	-----	AEAYRYADLLA	49
DB	1473	EREDEDLGP	AFVVSAGOR	WAASSVGL	FAGSNNIY	STSQSFVNTLD	SELFQ SARLSA	1532

QY 50 KENGYTADLEDGGYTINIRFA 71  
+ ||:||||: :||  
Db 1533 SSLRYYGLGLENGGYTTLQFA 1554

RESULT 11  
AC1498

C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1498

R; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloesch, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A; Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A.; Title: Comparative genomics of *Listeria* species.  
A; Reference number: AB1077; MUID: 21537279; PMID: 11679669  
A; Accession: AC1498







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:14:02 ; Search time 44.137 Seconds  
(without alignments)  
420.957 Million cell updates/sec

Title: US-08-325-278b-1\_COPY\_225\_296

Perfect score: 370

Sequence: 1 KEKTPPEPKKEVTIKANLIY.....GKYTADLEGGYTINIRFAG 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370	100.0	455	2 Q53291	Q53291 streptococ
2	370	100.0	719	2 Q51912	Q51912 peptostrept
3	329	88.9	992	2 Q51918	Q51918 peptostrept
4	80.5	21.8	1576	16 Q82EW8	Q82EW8 yersinia pe
5	74	20.0	324	3 Q9UUK3	Q9UUK3 schizosacch
6	74	20.0	846	2 Q8RPV1	Q8RPV1 streptococ
7	71.5	19.3	216	4 Q9BPV7	Q9BPV7 homo sapien
8	71.5	19.3	398	4 Q9BPZ1	Q9BPZ1 homo sapien
9	71.5	19.3	657	16 Q82DS2	Q82DS2 listeria in
10	71	19.2	871	10 Q9LME2	Q9LME2 arabidopsis
11	70.5	19.1	256	5 Q27039	Q27039 theileria p
12	70.5	19.1	256	5 Q27039	Q27039 theileria p
13	70.5	19.1	280	5 Q27030	Q27030 theileria p
14	70.5	19.1	398	11 Q9CZ02	Q9CZ02 mus musculu
15	70.5	19.1	398	11 Q91WD1	Q91WD1 mus musculu
16	70	18.9	549	17 Q8TZL6	Q8TZL6 pyrococcus

17	69.5	18.8	1086	10 Q9SGU0	Q9sgu0 arabidopsis
18	69.5	18.8	1849	2 Q9S4K2	Q9s4k2 lactobacill
19	69.5	18.8	2062	10 Q9C7J2	Q9c7j2 arabidopsis
20	69	18.6	397	16 Q92ED7	Q92ed7 listeria in
21	69	18.6	401	16 Q92ED5	Q92ed5 listeria in
22	69	18.6	1025	10 Q9LME3	Q9lme3 arabidopsis
23	68.5	18.5	858	10 Q9SGT8	Q9sgt8 arabidopsis
24	67.5	18.2	383	5 Q9VTV4	Q9vtv4 drosophila
25	67	18.1	278	10 Q94KL8	Q94kl8 podophyllum
26	67	18.1	283	3 Q03337	Q03337 saccharomyc
27	66.5	18.0	1029	10 Q9SGT9	Q9sgt9 arabidopsis
28	66	17.8	495	16 Q9W9G8	Q9w9g8 staphylococ
29	65.5	17.7	225	2 Q9WX49	Q9wx49 onion yello
30	65	17.7	4545	2 Q9X4W2	Q9x4w2 vibrio chol
31	65.5	17.7	4558	16 Q9KS12	Q9ks12 vibrio chol
32	65	17.6	304	6 Q9BE26	Q9be26 macaca fasc
33	65	17.6	319	4 Q00477	Q00477 homo sapien
34	65	17.6	334	4 Q9NR44	Q9nr44 homo sapien
35	65	17.6	334	4 Q9BU81	Q9bu81 homo sapien
36	65	17.6	357	4 Q15338	Q15338 homo sapien
37	65	17.6	359	4 P78410	P78410 homo sapien
38	65	17.6	495	4 Q9HCY1	Q9hcy1 homo sapien
39	65	17.6	513	4 Q00481	Q00481 homo sapien
40	65	17.6	584	4 Q00478	Q00478 homo sapien
41	64.5	17.4	279	2 Q9ADV8	Q9adv8 ehrlichia c
42	64	17.3	500	16 Q8NY41	Q8ny41 staphylococ
43	64	17.3	585	16 Q8EJS9	Q8ejs9 shewanella
44	64	17.3	1433	2 Q45616	Q45616 bacillus su
45	63.5	17.2	183	16 Q8EBW9	Q8ebw9 shewanella

#### ALIGNMENTS

##### RESULT 1

Q53291 ID Q53291 PRELIMINARY; PRT; 455 AA.  
AC Q53291;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Protein LG (Fragment).  
OS Streptococcus sp.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93094283; PubMed=1460053;  
RA Kihlberg B.M., Sjobring U., Kastern W., Bjorck L.;  
RT "Protein LG: a hybrid molecule with unique immunoglobulin binding properties."  
RL J. Biol. Chem. 267:25583-25588(1992).  
DR EMBL; S50809; AAA03280.1; -.  
DR HSSP; P06654; LPGX.  
DR InterPro; IPR003147; B1.  
DR InterPro; IPR000724; IgG\_bind\_B.  
DR Pfam; PF02246; B1; 4.  
DR Pfam; PF01378; IgG-binding\_B; 2.  
FT NON\_TER 455 455  
SQ SEQUENCE 455 AA; 49926 MW; 381FC235BBC8307B CRC64;

Query Match 100.0%; Score 370; DB 2; Length 455;  
Best Local Similarity 100.0%; Pred. No. 6.5e-30;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENKYYTADLE 60  
|||||  
Db 246 KEKTPPEPKKEVTIKANLIYADGKTQTAEFKGTFAEATAEAYRYADLLAKENKYYTADLE 305  
QY 61 DGGYTINIRFAG 72  
|||||  
Db 306 DGGYTINIRFAG 317

```
RESULT 2
Q51912
ID Q51912 PRELIMINARY; PRT; 719 AA.
AC Q51912;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=90215984; PubMed=2108927;
RA Kastern W., Holst E., Nielsen E., Sjobring U., Bjorck L.;
RT "Protein L, a bacterial immunoglobulin-binding protein and possible
virulence determinant.";
RL Infect. Immun. 58:1217-1222(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=92316971; PubMed=1618782;
RA Bjoerck L., Sjoerbing U., Kastern W.;
RT "Structure of peptostreptococcal protein L and identification of
repeated immunoglobulin light chain-binding domain.";
RL J. Biol. Chem. 267:12820-12825(1992).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
DR EMBL; M86697; AAA25612.1; -.
DR InterPro; IPR003147; B1.
DR ProDom; PD153432; Csurface_antigen; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 719 PROTEIN L.
SQ SEQUENCE 719 AA; 78983 MW; 963A8D76D5E34DD2 CRC64;

Query Match 100.0%; Score 370; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTEPEKKEVTIKANLIYADGKTQTAEFGTFAEATAYRYADLLAKENGYKTADLE 60
Db 318 KEKTEPEKKEVTIKANLIYADGKTQTAEFGTFAEATAYRYADLLAKENGYKTADLE 377

QY 61 DGGYTINIRFAG 72
Db 378 DGGYTINIRFAG 389

RESULT 3
Q51918
ID Q51918 PRELIMINARY; PRT; 992 AA.
AC Q51918;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
Perna N.T., Rose D.J., Mau B., Zhou S., Zhuker S., Schwartz D.C.,
Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=3316;
RX MEDLINE=95078460; PubMed=7987012;
RA Murphy J.P., Trowern A.R., Dugleby C.J.;
RT "Nucleotide sequence of the gene for peptostreptococcal protein L.";
RL DNA Seq. 4:259-265(1994).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
DR EMBL; L04466; AAA67503.1; -.
DR HSSP; Q51911; IGAB.
DR InterPro; IPR003147; B1.
DR InterPro; IPR002988; GA.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR Pfam; PF02246; B1; 4.
DR Pfam; PF01468; GA; 4.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 992 PROTEIN L.
SQ SEQUENCE 992 AA; 108700 MW; 9CFF5771578A5DCE CRC64;

Query Match 88.9%; Score 329; DB 2; Length 992;
Best Local Similarity 90.0%; Pred. No. 2.7e-25;
Matches 63; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 KTEPEKKEVTIKANLIYADGKTQTAEFGTFAEATAYRYADLLAKENGYKTADLE 62
Db 468 KTEPEKKEVTIKANLIYADGKTQTAEFGTFAEATAYRYADLLAKENGYKTADLE 527

QY 63 GYTINIRFAG 72
Db 528 GYTINIRFAG 537

RESULT 4
Q8ZEVS
ID Q8ZEVS PRELIMINARY; PRT; 1576 AA.
AC Q8ZEVS;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Putative hemolysin.
GN YPO2045 OR Y2267.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
Perna N.T., Rose D.J., Mau B., Zhou S., Zhuker S., Schwartz D.C.,
Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
```

RA Levesque C., Vadeboncoeur C., Frenette M.;  
RT "Isolation and Characterization of a Tn917 Insertion Fimbria-Deficient  
RT Mutant of Streptococcus salivarius: Identification of a Regulatory  
RT Protein Controlling the Expression of Fimbriae and Other Cell Surface  
RT Proteins";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY  
CC AN AMIDE BOND (BY SIMILARITY).  
DR EMBL; AF353638; AAL83961.1; -  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR006192; LPXTG.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor.  
SQ SEQUENCE 846 AA; 90589 MW; D02F33F4B71EC0E9 CRC64;

Query Match 20.0%; Score 74; DB 2; Length 846;  
Best Local Similarity 36.6%; Pred. No. 42;  
Matches 26; Conservative 10; Mismatches 25; Indels 10; Gaps 5;

QY 3 KTEPEEPKEEVTIKANLIYADGKQTQTAEFGTFAEATAEAYRYADLLAKENGYTA-- 57  
Db 604 KTPTAPKPHKSVYDSKNDI--DGKT-TDDDKVIF-HLTDTSPYKDMAASKKALEFALL 659

QY 58 -DLEDGGYTN 67  
Db 660 DDVQDGYATVD 670

RESULT 7  
Q9BPV7 PRELIMINARY; PRT; 216 AA.  
ID Q9BPV7  
AC Q9BPV7;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Similar to BNS1 (BHK21) temperature sensitivity complementing.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC003039; AAH03039.1; -  
DR EMBL; BC000516; AAH00516.1; -  
SQ SEQUENCE 216 AA; 24806 MW; F29028AEACA8DB04 CRC64;

Query Match 19.3%; Score 71.5; DB 4; Length 216;  
Best Local Similarity 29.6%; Pred. No. 15;  
Matches 24; Conservative 9; Mismatches 29; Indels 19; Gaps 3;

QY 3 KTEPEEPKEEVTIKANLIYAD-----GKQTQTAEFGTFAEATAEAYRYADLLAK 50  
Db 66 KIKEEPKEEVTIVKKEKREDRDRQREHGHRGRPREVTQSHSIFEQGPAAEMK-----KK 120

QY 51 ENKGYTADLEDGG--YTINIR 69  
Db 121 GNWDKTVDSVDMGSPSHIINIK 141

RESULT 8  
Q9BPZ1 PRELIMINARY; PRT; 398 AA.  
ID Q9BPZ1  
AC Q9BPZ1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Similar to BNS1 (BHK21) temperature sensitivity complementing.  
OS Homo sapiens (Human).  
RP SEQUENCE FROM N.A.

RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL; AJ414151; CAC90857.1; -  
DR EMBL; AE013829; AAM85827.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 1576 AA; 169976 MW; 7029EC4BA386424C CRC64;

Query Match 21.8%; Score 80.5; DB 16; Length 1576;  
Best Local Similarity 25.7%; Pred. No. 19;  
Matches 28; Conservative 13; Mismatches 25; Indels 43; Gaps 4;

QY 1 KETPEEPKEEVTIKANLIYADGKQTQTAEFG-----TFAE----- 36  
Db 1001 KANTPEQKEGVSIRG-----GWTATQEIKGHLGVKRAETSGQDSYAEMLVGNNAKSG 1054

QY 37 ----ATAEAYRYADLLAKENGYTAD-----LEDGGYTNIRFAG 72  
Db 1055 VSIKTTGDAYYATNIEGGNGDITDAGNNLYFDQVDQSQRSSNIKFSG 1103

RESULT 5  
Q9UUK3 PRELIMINARY; PRT; 324 AA.  
ID Q9UUK3  
AC Q9UUK3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical coiled-coil protein.  
GN SPAC1952.03.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RA McDougall R.C., Rajandream M.A., Barrell B.G., Bothe G., Pohl T.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL109820; CAB52567.1; -  
DR GeneDB\_SPombe; SPAC1952.03; -  
DR InterPro; IPR003323; OTU.  
DR Pfam; PF02338; OTU; 1.  
DR PROSITE; PS50802; OTU; 1.  
SQ SEQUENCE 324 AA; 37516 MW; 48BBB435960D4B9C CRC64;

Query Match 20.0%; Score 74; DB 3; Length 324;  
Best Local Similarity 30.2%; Pred. No. 13;  
Matches 19; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 KETPEEPKEEVTIKANLIYADGKQTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
Db 126 KENTPOQPKSRNRQKERL----ERRKAEKMKSSQAELESEKMAADLKNEEKFKFSKILE 181

QY 61 DGS 63  
Db 182 EAG 184

RESULT 6  
Q8RPV1 PRELIMINARY; PRT; 846 AA.  
ID Q8RPV1  
AC Q8RPV1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Cell surface protein B.  
GN CSFB.  
OS Streptococcus salivarius.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1304;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Levesque C., Vadeboncoeur C., Frenette M.;  
RT "Isolation and Characterization of a Tn917 Insertion Fimbria-Deficient  
RT Mutant of Streptococcus salivarius: Identification of a Regulatory  
RT Protein Controlling the Expression of Fimbriae and Other Cell Surface  
RT Proteins";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
CC -|- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY  
CC AN AMIDE BOND (BY SIMILARITY).  
DR EMBL; AF353638; AAL83961.1; -  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR006192; LPXTG.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR TIGRFAMs; TIGR01167; LPXTG\_anchor; 1.  
DR PROSITE; PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Cell wall; Peptidoglycan-anchor.  
SQ SEQUENCE 846 AA; 90589 MW; D02F33F4B71EC0E9 CRC64;

Query Match 20.0%; Score 74; DB 2; Length 846;  
Best Local Similarity 36.6%; Pred. No. 42;  
Matches 26; Conservative 10; Mismatches 25; Indels 10; Gaps 5;

QY 3 KTPEEPKEEVTIKANLIYADGKQTQTAEFGTFAEATAEAYRYADLLAKENGYTA-- 57  
Db 604 KTPTAPKPHKSVYDSKNDI--DGKT-TDDDKVIF-HLTDTSPYKDMAASKKALEFALL 659

QY 58 -DLEDGGYTN 67  
Db 660 DDVQDGYATVD 670

RESULT 7  
Q9BPV7 PRELIMINARY; PRT; 216 AA.  
ID Q9BPV7  
AC Q9BPV7;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Similar to BNS1 (BHK21) temperature sensitivity complementing.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC003039; AAH03039.1; -  
DR EMBL; BC000516; AAH00516.1; -  
SQ SEQUENCE 216 AA; 24806 MW; F29028AEACA8DB04 CRC64;

Query Match 19.3%; Score 71.5; DB 4; Length 216;  
Best Local Similarity 29.6%; Pred. No. 15;  
Matches 24; Conservative 9; Mismatches 29; Indels 19; Gaps 3;

QY 3 KTPEEPKEEVTIKANLIYAD-----GKQTQTAEFGTFAEATAEAYRYADLLAK 50  
Db 66 KIKEEPKEEVTIVKKEKREDRDRQREHGHRGRPREVTQSHSIFEQGPAMMK-----KK 120

QY 51 ENKGYTADLEDGG--YTINIR 69  
Db 121 GNWDKTVDSMDGSPSHIINIK 141

RESULT 8  
Q9BPZ1 PRELIMINARY; PRT; 398 AA.  
ID Q9BPZ1  
AC Q9BPZ1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Similar to BNS1 (BHK21) temperature sensitivity complementing.  
OS Homo sapiens (Human).  
RP SEQUENCE FROM N.A.

RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL; AJ414151; CAC90857.1; -  
DR EMBL; AE013829; AAM85827.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 1576 AA; 169976 MW; 7029EC4BA386424C CRC64;

Query Match 21.8%; Score 80.5; DB 16; Length 1576;  
Best Local Similarity 25.7%; Pred. No. 19;  
Matches 28; Conservative 13; Mismatches 25; Indels 43; Gaps 4;

QY 1 KEKTEEPKEEVTIKANLIYADGKQTQTAEFG-----TFAE----- 36  
Db 1001 KANTPEQKEGVSIRG-----GWTATQEIKGHLGVKRAETSGQDSYAEMLVGNNAKSG 1054

QY 37 ----ATAEAYRYADLLAKENGYTAD-----LEDGGYTNIRFAG 72  
Db 1055 VSIKTTGDAYYATNIEGGNGDITDAGNNLYFDQVDQSQRSSNIKFSG 1103

RESULT 5  
Q9UUK3 PRELIMINARY; PRT; 324 AA.  
ID Q9UUK3  
AC Q9UUK3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical coiled-coil protein.  
GN SPAC1952.03.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RA McDougall R.C., Rajandream M.A., Barrell B.G., Bothe G., Pohl T.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL109820; CAB52567.1; -  
DR GeneDB\_SPombe; SPAC1952.03; -  
DR InterPro: IPR003323; OTU.  
DR Pfam; PF02338; OTU; 1.  
DR PROSITE; PS50802; OTU; 1.  
SQ SEQUENCE 324 AA; 37516 MW; 48BBB435960D4B9C CRC64;

Query Match 20.0%; Score 74; DB 3; Length 324;  
Best Local Similarity 30.2%; Pred. No. 13;  
Matches 19; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 KEKTEEPKEEVTIKANLIYADGKQTQTAEFGTFAEATAEAYRYADLLAKENGYTADLE 60  
Db 126 KENTPOQPKSRNRQKERL----ERRKAEKMKSSQAELESEKMDLKNEEKKFKSKILE 181

QY 61 DGS 63  
Db 182 EAG 184

RESULT 6  
Q8RPV1 PRELIMINARY; PRT; 846 AA.  
ID Q8RPV1  
AC Q8RPV1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Cell surface protein B.  
GN CSFB.  
OS Streptococcus salivarius.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1304;  
RN [1]  
RP SEQUENCE FROM N.A.

[illegible]



Search completed: September 3, 2003, 11:23:15  
Job time : 46.137 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 11:14:02; Search time 46.589 Seconds  
(without alignments)  
420.957 Million cell updates/sec

Title: US-08-325-278B-1\_COPY\_5\_80

Perfect score: 389

Sequence: 1 KEETPETPTDSEEEVITKA.....GEYTVADKGYTLNIKFAAG 76

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	100.0	455	2 Q53291	Q53291 streptococ
2	389	100.0	719	2 Q51912	Q51912 peptostrept
3	298.5	76.7	992	2 Q51918	Q51918 peptostrept
4	85	21.9	1576	16 Q82EV8	Q82EV8 yersinia pe
5	74	19.0	529	1 Q48937	Q48937 methanosarc
6	73.5	18.9	451	2 Q924J9	Q924J9 lactobacill
7	72	18.5	477	10 Q932G9	Q932G9 arabidopsis
8	72	18.5	487	10 Q9STK2	Q9STK2 arabidopsis
9	71.5	18.4	947	2 Q86487	Q86487 staphylococ
10	71	18.3	528	16 Q9KX8	Q9KX8 bacillus ha
11	71	18.3	929	3 P78718	P78718 nectria hae
12	70.5	18.1	955	16 Q8NXX7	Q8NXX7 staphylococ
13	70	18.0	645	5 Q3862	Q3862 ascaris suu
14	69.5	17.9	292	16 Q92FA8	Q92FA8 listeria in
15	69.5	17.9	300	11 Q9JHY1	Q9JHY1 rattus norv
16	69	17.7	190	17 Q8TQF1	Q8TQF1 methanosarc

17	69	17.7	341	3 Q12191	Q12191 saccharomyc
18	69	17.7	1430	5 Q9VMA7	Q9VMA7 drosophila
19	68.5	17.6	401	16 Q8FHX3	Q8FHX3 escherichia
20	68	17.5	574	2 Q9AJD4	Q9AJD4 streptococ
21	67.5	17.4	218	2 Q9RF14	Q9RF14 streptococ
22	67.5	17.4	1022	2 Q93T54	Q93T54 streptococ
23	67.5	17.4	1056	16 Q9PAQ0	Q9PAQ0 xyliella fas
24	67.5	17.4	1876	16 Q8DQNS	Q8DQNS streptococ
25	67	17.2	529	17 Q8TV18	Q8TV18 methanopyru
26	67	17.2	574	16 Q8P2T7	Q8P2T7 streptococ
27	67	17.2	588	2 Q9S0T6	Q9S0T6 escherichia
28	67	17.2	779	10 Q9SZB6	Q9SZB6 arabidopsis
29	67	17.2	890	2 Q01891	Q01891 enterococ
30	67	17.2	891	2 Q04111	Q04111 enterococ
31	67	17.2	1433	2 Q45616	Q45616 bacillus su
32	67	17.2	1881	16 Q8RGK2	Q8RGK2 fusobacteri
33	66.5	17.1	255	16 Q99XB4	Q99XB4 staphylococ
34	66.5	17.1	465	2 Q9X775	Q9X775 mycoplasma
35	66.5	17.1	642	16 Q9JTL4	Q9JTL4 neisseria m
36	66.5	17.1	843	2 Q47802	Q47802 enterococ
37	66	17.0	183	16 Q9KDD6	Q9KDD6 bacillus ha
38	66	17.0	489	17 Q8THX7	Q8THX7 methanosarc
39	66	17.0	498	16 Q8X6R2	Q8X6R2 escherichia
40	66	17.0	653	16 Q67097	Q67097 aquifex aso
41	66	17.0	1169	16 Q8K5Q0	Q8K5Q0 streptococ
42	66	17.0	1615	2 Q9KKA4	Q9KKA4 rickettsia
43	65.5	16.8	313	5 Q76641	Q76641 caenorhabdi
44	65.5	16.8	456	11 Q9CS77	Q9CS77 mus musculu
45	65.5	16.8	868	17 Q8PZM5	Q8PZM5 methanosarc

#### ALIGNMENTS

#### RESULT 1

Q53291 ID Q53291 PRELIMINARY; PRT; 455 AA.  
AC Q53291;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Protein LG (Fragment).  
OS Streptococcus sp.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93094283; PubMed=1460053;  
RA Kihlberg B.M., Sjöbring U., Kastern W., Björck L.;  
RT "Protein LG: a hybrid molecule with unique immunoglobulin binding  
properties";  
RL J. Biol. Chem. 267:25583-25588(1992).  
DR EMBL; S50809; AAA03280.1; -;  
DR HSSP; P06654; IPGX.  
DR InterPro; IPR003147; B1.  
DR InterPro; IPR000724; IgG\_bind\_B.  
DR Pfam; PF02246; B1; 4.  
DR Pfam; PF01378; IgG\_binding\_B; 2.  
FT NON\_TER 455 455  
SQ SEQUENCE 455 AA; 49926 MW; 381FC235B8C8307B CRC64;

Query Match 100.0%; Score 389; DB 2; Length 455;

Best Local Similarity 100.0%; Pred. No. 2.1e-30;

Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETPTDSEEEVITKANLIFANGSTQTAETFGTFKATSEAYAYADTLKKNGEYT 60

|||||

Db 26 KEETPETPTDSEEEVITKANLIFANGSTQTAETFGTFKATSEAYAYADTLKKNGEYT 85

|||||

QY 61 DVADKGYTLNIKFAAG 76

|||||

Db 86 DVADKGYTLNIKFAAG 101

```

RESULT 2
ID Q51912 PRELIMINARY; PRT; 719 AA.
AC Q51912;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=90215984; PubMed=2108927;
RA Kastern W., Holst E., Nielsen E., Sjoerding U., Bjorck L.;
RT "Protein L, a bacterial immunoglobulin-binding protein and possible
FT virulence determinant.";
RL Infect. Immun. 58:1217-1222(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=92316971; PubMed=1618782;
RA Bioerck L., Sjoerding U., Kastern W.;
RT "Structure of peptostreptococcal protein L and identification of
FT repeated immunoglobulin light chain-binding domain.";
RL J. Biol. Chem. 267:12820-12825(1992).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
EMBL: M86697; AAA25612.1; -
DR InterPro: IPR003147; B1.
DR InterPro: IPR004829; Csurface_antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 5.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR ProDom: PD153432; Csurface_antigen; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 719 PROTEIN L.
SQ SEQUENCE 719 AA; 78983 MW; 963A8D76D5E34DD2 CRC64;

Query Match 100.0%; Score 389; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.5e-30;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETDSEEVTKANLIFANGSTQTAEFKGTSEAYAYADTLKNDGEY 60
DB 98 KEETPETDSEEVTKANLIFANGSTQTAEFKGTSEAYAYADTLKNDGEY 157
QY 61 VDVAADKGYTLNKFAG 76
DB 158 VDVAADKGYTLNKFAG 173

RESULT 3
ID Q51918 PRELIMINARY; PRT; 992 AA.
AC Q51918;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=90215984; PubMed=2108927;
RA Kastern W., Holst E., Nielsen E., Sjoerding U., Bjorck L.;
RT "Protein L, a bacterial immunoglobulin-binding protein and possible
FT virulence determinant.";
RL Infect. Immun. 58:1217-1222(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=312;
RX MEDLINE=92316971; PubMed=1618782;
RA Bioerck L., Sjoerding U., Kastern W.;
RT "Structure of peptostreptococcal protein L and identification of
FT repeated immunoglobulin light chain-binding domain.";
RL J. Biol. Chem. 267:12820-12825(1992).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
EMBL: M86697; AAA25612.1; -
DR InterPro: IPR003147; B1.
DR InterPro: IPR004829; Csurface_antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 5.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR ProDom: PD153432; Csurface_antigen; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 719 PROTEIN L.
SQ SEQUENCE 719 AA; 78983 MW; 963A8D76D5E34DD2 CRC64;

Query Match 100.0%; Score 389; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.5e-30;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEETPETDSEEVTKANLIFANGSTQTAEFKGTSEAYAYADTLKNDGEY 60
DB 98 KEETPETDSEEVTKANLIFANGSTQTAEFKGTSEAYAYADTLKNDGEY 157
QY 61 VDVAADKGYTLNKFAG 76
DB 158 VDVAADKGYTLNKFAG 173

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=3316;
RX MEDLINE=95078460; PubMed=7987012;
RA Murphy J.P., Tricern A.R., Duggieby C.J.;
RT "Nucleotide sequence of the gene for peptostreptococcal protein L.";
RL DNA Seq. 4:259-265(1994).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
EMBL: L04466; AAA67503.1; -
DR HSP; Q51911; IGAB.
DR InterPro: IPR003147; B1.
DR InterPro: IPR002988; GA.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; B1; 4.
DR Pfam: PF01468; GA; 4.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 992 PROTEIN L.
SQ SEQUENCE 992 AA; 108700 MW; 9CFE5771578A5DCE CRC64;

Query Match 76.7%; Score 298.5; DB 2; Length 992;
Best Local Similarity 80.3%; Pred. No. 4.7e-21;
Matches 61; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 1 KEETPETDSEEVTKANLIFANGSTQTAEFKGTSEAYAYADTLKNDGEY 60
DB 247 EKETPE-----PEEVTIKANLIFADGSTQNAEFKGTFAKAVSDAYADALAKNDGEY 301
QY 61 VDVAADKGYTLNKFAG 76
DB 302 VDVAADKGYTLNKFAG 317

RESULT 4
ID Q8ZEVE8 PRELIMINARY; PRT; 1576 AA.
AC Q8ZEVE8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative hemolysin.
GN YPO2045 OR Y2267.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Zhong S., R. R., Plano G.V.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Straley S.C.,
RA Perry R.D.;
RT "Genome sequence of Versinia pestis KIM.";
RN [1]

```

```

RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AJ414151; CAC90857.1; -.
DR EMBL; AE013829; AAM85827.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1576 AA; 169976 MW; 7029EC4BA386424C CRC64;

Query Match      21.9%; Score 85; DB 16; Length 1576;
Best Local Similarity 31.2%; Pred. No. 11;
Matches 25; Conservative 11; Mismatches 26; Indels 18; Gaps 3;

QY 6 ETPETSEEVTKANLIFANGSTOTAEFKGTSEAYAYADTLKKDNGEYTDV--- 62
Db 1033 ETSQGSYAEMLVGNL---NAKS-----GVSIKTTGDYYATNIEGGNGDITIDAGN 1083
QY 63 -----VADKGYTLNKFAG 76
Db 1084 NLYFDQVQDSORSSNKFSG 1103

RESULT 5
Q48937 PRELIMINARY; PRT; 529 AA.
AC Q48937;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Orf4 protein (Fragment).
GN ORF4.
OS Methanosarcina barkeri.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2208;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fusaro/DSM 804;
RX MEDLINE=96184912; PubMed=8617280;
RA Vorholt J.A., Vaupel M., Thauer R.K.;
RT "A polyferredoxin with eight [4Fe-4S] clusters as a subunit of
RT molybdenum formylmethanofuran dehydrogenase from Methanosarcina
RT barkeri.";
RL Eur. J. Biochem. 236:309-317(1996).
DR EMBL; X93084; CAA63625.1; -.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5; 1.
FT NON_TER 529 529
SQ SEQUENCE 529 AA; 58606 MW; 218001E917C0F7F4 CRC64;

Query Match      19.0%; Score 74; DB 1; Length 529;
Best Local Similarity 38.3%; Pred. No. 37;
Matches 23; Conservative 8; Mismatches 23; Indels 6; Gaps 3;

QY 15 EVTKANLIFANGSTOTAEFKGTSEAYAYADTLKKDNGEYTDVADKGYTLNIK 73
Db 119 EFKLRNNVTHDGSKMTAEVDNTEKLVISENAKVASMLKID---SIEIVD-NYTLKIK 173

RESULT 6
Q924J9 PRELIMINARY; PRT; 451 AA.
AC Q924J9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE S-layer protein precursor.
GN S-LAYER.
OS Lactobacillus crispatus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=47770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M247;
RA Ventura M., Callegari M., Morelli L.;

```

```

RT "Detection of a strain specific region in the S-layer sequence of
RT Lactobacillus crispatus.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ007839; CAA07708.1; -.
DR InterPro; IPR004903; SLAP.
DR Pfam; PF03217; SLAP; 1.
KW S-layer; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 451 S-LAYER PROTEIN.
SQ SEQUENCE 451 AA; 47012 MW; 7F79C6D0F7C458F7 CRC64;

Query Match      18.9%; Score 73.5; DB 2; Length 451;
Best Local Similarity 35.8%; Pred. No. 35;
Matches 24; Conservative 3; Mismatches 15; Indels 25; Gaps 3;

QY 26 NGSTQTAERFKGTSEAYAYADTLKKDNGEYTDV-----ADKGYTL----- 70
Db 107 NGKVSVDKSGT-----AVTDFSKLTNGSYTTVSGVSNFGTANKTITLGSKN 157
QY 71 -NIKFGAG 76
Db 158 SNVFKFAG 164

RESULT 7
Q93ZG9 PRELIMINARY; PRT; 477 AA.
AC Q93ZG9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AT4925340/T30C3_20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057543; AAL09783.1; -.
DR EMBL; AY143978; AAN28917.1; -.
DR InterPro; IPR001179; FKBP_PPIase.
DR Pfam; PF00254; FKBP; 1.
DR PROSITE; PS00454; FKBP_PPIASE_2; 1.
DR PROSITE; PSS0059; FKBP_PPIASE_3; 1.
SQ SEQUENCE 477 AA; 52223 MW; 10C405953B17675E CRC64;

Query Match      18.5%; Score 72; DB 10; Length 477;
Best Local Similarity 26.4%; Pred. No. 52;
Matches 23; Conservative 17; Mismatches 27; Indels 20; Gaps 4;

```

```

QY 3 ETPETETSEEEVITKANLIFANGSTQTAETAEFGKTEKAT-----SEAYAYAD----- 50
Db 319 ESSKTPDKSAEKKTKKKK---KKPSDEAAEISGTVKEKQTPADSKSSQVRTYTPNGLIVEE 375
QY 51 -TLKDNNGEYTDVADKGYTLNKKFAG 76
Db 376 LSMGKPNGRK----ADPGKTVSVRYIG 398

RESULT 8
Q9STK2 PRELIMINARY; PRT; 487 AA.
AC Q9STK2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical 53.3 kDa protein.
GN T30C3.20 OR AT4G25340.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eusoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA Meyer M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA May K.F.X., Lemcke K., Schueller C.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL079350; CAB45512.1; -.
DR EMBL; AL161563; CAB81345.1; -.
DR HSSP; O00688; 1PBK.
DR InterPro; IPR001179; FKBP_PP1ase.
DR Pfam; PF00254; FKBP; 1.
DR PROSITE; PS00454; FKBP_PP1ase_2; 1.
DR PROSITE; PS50059; FKBP_PP1ase_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 487 AA; 53290 MW; 5B2FA21570BC0AC6 CRC64;

Query Match 18.5%; Score 72; DB 10; Length 487;
Best Local Similarity 26.4%; Pred. No. 54;
Matches 23; Conservative 17; Mismatches 27; Indels 20; Gaps 4;

QY 3 ETPETETSEEEVITKANLIFANGSTQTAETAEFGKTEKAT-----SEAYAYAD----- 50
Db 329 ESSKTPDKSAEKKTKKKK---KKPSDEAAEISGTVKEKQTPADSKSSQVRTYTPNGLIVEE 385
QY 51 -TLKDNNGEYTDVADKGYTLNKKFAG 76
Db 386 LSMGKPNGRK----ADPGKTVSVRYIG 408

RESULT 9
O86487 PRELIMINARY; PRT; 947 AA.
AC O86487;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SdrC protein.
GN "SDRC.
OS Staphylococcus aureus.

```

```

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=9098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
DR EMBL; AJ005645; CA006650.1; -.
DR InterPro; IPR005877; Gpos_YSIrk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIrk_signal; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIrk_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 947 AA; 102888 MW; 3C6EFD6E35121554 CRC64;

Query Match 18.4%; Score 71.5; DB 2; Length 947;
Best Local Similarity 30.6%; Pred. No. 1.3e+02;
Matches 22; Conservative 17; Mismatches 22; Indels 11; Gaps 5;

QY 7 TPETDSEEEVITKANLIFAN-GSTQTAEEFKGTFEKATSEAY-----AYADTLKKDNG-- 57
Db 407 TPETSKLDVTDQFDVIYNDNKNATVDLMKG--QTSSNKKYIIQQVAYPDNSSTDNGKI 464

QY 58 EYTDVDVADKGYT 69
Db 465 DYTLDTDTKTKYS 476

RESULT 10
Q9KBX8 PRELIMINARY; PRT; 528 AA.
AC Q9KBX8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Nickel ABC transporter (Nickel-binding protein).
GN BHL796.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AF001513; BAB05515.1; -.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5; 1.
KW Complete proteome.
SQ SEQUENCE 528 AA; 59381 MW; BC709550C21A66B2 CRC64;

Query Match 18.3%; Score 71; DB 16; Length 528;
Best Local Similarity 30.8%; Pred. No. 74;
Matches 20; Conservative 13; Mismatches 24; Indels 8; Gaps 3;

QY 10 TDSEEEVITKANLIFANGSTQTAEEFKGTFEKATSEAYAYADTLKKDNGEYTDVADK 67
Db 106 TDGQHWTKLRDVTDFQNGKEMDAEAVKSLERALDEVAIENALKID-----ETADG 159

```

```

QY 68 YTLNI 72
   |||:|
Db 160 YTLHI 164

RESULT 11
P78718
ID P78718 PRELIMINARY; PRT; 929 AA.
AC P78718;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Kinesin.
GN NHKIN1.
OS Nectria haematococca.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Nectriaceae; Nectria.
OX NCBI_TaxID=140110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T213 mating population VI;
RA Aist J.R., Wu Q., Wiersel S.G., Turgeon B.G., Yoder O.C.;
RT "Disruption of a fungal kinesin gene affects both morphogenesis and
RT mitosis.";
RL Mol. Biol. Cell 0:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=T213 mating population VI;
RA Wu Q., Aist J.R., Wiersel S.G., Turgeon B.G., Yoder O.C., Sandrock T.;
RT "Nectria haematococca mating population VI NHRAB6 and Nkln1 genes.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U86521; AAB47851.1;
DR HSSP; P33176; I8G2.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00639; THIOLEPROTEASE_HIS; 1.
DR ATP-binding; Coiled coil; Microtubules; Motor protein.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 929 AA; 103226 MW; 37C1A9CAF6003608 CRC64;

Query Match 18.3%; Score 71; DB 3; Length 929;
Best Local Similarity 32.1%; Pred. No. 1.4e+02;
Matches 27; Conservative 11; Mismatches 26; Indels 20; Gaps 4;

QY 1 KEETPTPTDSE-----EVTIKANLIFANGSTQTAEKGFTEKAT----SEAYAYADTL 52
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 469 KEELYLKEDHSKVDKENEKLTVEV-----EFKQLERLTFESKEAQITMDTL 517
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 53 KKDNGEYTVDVAD-KGYTLNLIKFA 75
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 518 KEANTLTLTDLDDVKQQLLDVKMS 541
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 12
Q8NXX7
ID Q8NXX7 PRELIMINARY; PRT; 955 AA.
AC Q8NXX7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser-Asp rich fibrinogen-binding bone sialoprotein-binding
DE protein.
GN SDRG OR MW0516.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AF004823; BAB94381.1;
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 955 AA; 103642 MW; 65FD8BB93A477A22 CRC64;

Query Match 18.1%; Score 70.5; DB 16; Length 955;
Best Local Similarity 28.2%; Pred. No. 1.7e+02;
Matches 20; Conservative 16; Mismatches 26; Indels 9; Gaps 4;

QY 7 TPETSEBEVTIKANLIFAN-GSTQTAEKGFTEKATSEAY-----AYADTLKKONG--E 58
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 407 TPDTSKLKDVTFNQFNITYSNDNKTATVDLMNG-OTSSNKQYIIQVAYPDNTSTONGKID 465
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 59 YTVDVADKGYT 69
   |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 466 YTLDTDKTKYS 476

RESULT 13
Q33862
ID Q33862 PRELIMINARY; PRT; 645 AA.
AC Q33862;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Flavoprotein subunit of complex II precursor.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95258075; PubMed=7739664;
RA Kuramochi T., Hirawake H., Kojima S., Takamiya S., Furushima R.,
RA Aoki T., Komuniecki R., Kita K.;
RT "Sequence comparison between the flavoprotein subunit of the fumarate
RT reductase (ComplexII) of the anaerobic parasitic nematode, Ascaris
RT suum and the succinate dehydrogenase of the aerobic, free-living
RT nematode, Caenorhabditis elegans.";
RL Mol. Biochem. Parasitol. 68:177-187(1994).
DR EMBL; D30650; BAA21636.1;
DR HSSP; P00363; IKF6.
DR InterPro; IPR003953; FAD_bind2.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR003952; FRD/SDH_FAD.
DR InterPro; IPR00205; NAD_binding.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR004112; Succ_DH_flav_C.
DR Pfam; PF00890; FAD_binding_2; 1.
DR Pfam; PF02910; succ_DH_flav_C; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASEI.
DR PROSITE; PS00504; FRD_SDH_FAD_BINDING; 1.
KW Transit peptide.
FT TRANSIT 1 30 POTENTIAL.
FT CHAIN 31 645 FLAVOPROTEIN SUBUNIT OF COMPLEX II.
SQ SEQUENCE 645 AA; 71201 MW; 0F894C36D3B2D68E CRC64;

Query Match 18.0%; Score 70; DB 5; Length 645;

```



Result No.	Score	Query Match	Length	DB	ID	Description
1	371	100.0	455	2	Q53291	streptococc
2	371	100.0	719	2	Q51912	peptostrept
3	308	83.0	992	2	Q51918	peptostrept
4	98.5	26.5	216	4	Q9BPV7	homo sapien
5	98.5	26.5	398	4	Q9BPZ1	homo sapien
6	93.5	25.2	398	11	Q9C202	mus musculu
7	93.5	25.2	398	11	Q91WD1	mus musculu
8	89.5	24.1	1576	16	Q8ZEV8	yersinia p
9	73	19.7	588	2	Q9S0T6	escherichia
10	70	18.9	1498	5	Q8IDP2	plasmodium
11	69.5	18.7	256	5	Q27039	theileria p
12	69.5	18.7	256	5	Q27029	theileria p
13	69.5	18.7	279	2	Q9ADV8	ehrlichia c
14	69.5	18.7	280	5	Q27030	theileria p
15	69.5	18.7	451	2	Q9Z4J9	lactobacill
16	69.5	18.7	623	2	Q9F4L0	fibrobacter

```

RESULT 2
Q51912
ID Q51912 PRELIMINARY; PRT; 719 AA.
AC Q51912;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]
SEQUENCE FROM N.A.
RP RP
RC STRAIN=312;
RX MEDLINE=90215984; PubMed=2108927;
RA Kastern W., Holst E., Nielsen E., Sjoerling U., Bjorck L.;
RT "protein L, a bacterial immunoglobulin-binding protein and possible
RT virulence determinant";
RL Repeated immunoglobulin light chain-binding domain.";
RC J. Biol. Chem. 267:12820-12825(1992).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
CC AN AMIDE BOND (BY SIMILARITY).
CC EMBL; M86697; AAA25612.1; -.
DR InterPro: IPR003147; BL.
DR InterPro: IPR004829; Csurface_antigen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR006192; LPXTG.
DR Pfam: PF02246; BL; 5.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR ProDom: PD153432; Csurface_antigen; 1.
DR TIGRfams: TIGR01167; LPXTG_anchor; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 18
FT CHAIN 19 719
FT PROTEIN L.
SQ SEQUENCE 719 AA; 78983 MW; 963A8D76D5E34DD2 CRC64;
Query Match 100.0%; Score 371; DB 2; Length 719;
Best Local Similarity 100.0%; Pred No. 8.1e-29;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGFEEATAEAYRYADALKKDNGEYTVDDVA 60
Db 174 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGFEEATAEAYRYADALKKDNGEYTVDDVA 233
QY 61 DKGYTLNIRKFAF 72
Db 234 DKGYTLNIRKFAF 245

RESULT 3
Q51918
ID Q51918 PRELIMINARY; PRT; 992 AA.
AC Q51918;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein L precursor.
OS Peptostreptococcus magnus.
OC Bacteria; Firmicutes; Clostridia; Clostridiales;
OC Peptostreptococcaceae; Finegoldia.
OX NCBI_TaxID=1260;
RN [1]

```



DA  
RQD, MGT: 1014315, 2010420M171R.  
SQ SEQUENCE 398 AA; 44356 MW; B70179A124C1231B CRC64;

```

RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL: AJ414151; CAC90857.1; -.
DR EMBL: AB013829; AAM85827.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1576 AA; 169976 MW; 7029EC4BA386424C CRC64;

Query Match      24.1%; Score 89.5; DB 16; Length 1576;
Best Local Similarity 27.5%; Pred. No. 3.4;
Matches 30; Conservative 11; Mismatches 25; Indels 43; Gaps 4;

QY 1 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGTF-----EE----- 36
Db 1001 KANTTQEKGEVSLRG-----GMTATQEIKGHLGKVAETSGQDSVAEMLVGNINAKSG 1054

QY 37 ----ATAEAYRYADALKDNGEYTVD-----VADKGYTLNKFAG 72
Db 1055 VSIKTTGDYYYATNIEGGNDITTDAGNNLYFDQVDQSRSNKFSG 1103

RESULT 9
Q9S0T6 PRELIMINARY; PRT; 588 AA.
ID Q9S0T6
AC Q9S0T6
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Flagellin.
GN FLIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K42;
RA Ohnishi K., Ishioka K., Matsuba T., Harayama S.;
RT "Cloning of H antigen genes in E.coli serotypes and expression in
RT E.coli K-12.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB028476; BAA85085.1; -.
DR InterPro: IPR001492; Flagellin.
DR InterPro: IPR001029; Flagellin_C.
DR Pfam: PF00700; Flagellin_C; 1.
DR Pfam: PF00669; Flagellin_N; 1.
DR PRINTS: PR00207; FLAGELLIN.
DR PRODOM: PD000316; FLAGELLIN_C; 2.
SQ SEQUENCE 588 AA; 60458 MW; 9FFFE6D819DFE55D CRC64;

Query Match      19.7%; Score 73; DB 2; Length 588;
Best Local Similarity 24.6%; Pred. No. 48;
Matches 29; Conservative 19; Mismatches 20; Indels 50; Gaps 7;

QY 5 PPEPKEEV-----TKANLIYAD---GKTQTAEF-----KGTFEATA-EAYRY 44
Db 367 PDDASDDVLGTISYSKSKVDVLADTKATGNTTTFDNFNSGIMTSKVSFDAGTSTDTFKD 426

QY 45 AD-----ALKKDNGEYF-----VDVADK-----GYTLNKFAG 72
Db 427 ADGALTKEYTTSYAVNKDGEVTVADYAAVDSADKAVDDTKYKPTIGATVNLNSAG 484

RESULT 10

```

```

Q8IDP2 PRELIMINARY; PRT; 1498 AA.
ID Q8IDP2
AC Q8IDP2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN MALI3P1.226.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL844509; CAD52577.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1498 AA; 177555 MW; 10B973D7FE70F301 CRC64;

Query Match      18.9%; Score 70; DB 5; Length 1498;
Best Local Similarity 28.7%; Pred. No. 2.9e+02;
Matches 25; Conservative 14; Mismatches 28; Indels 20; Gaps 4;

QY 1 KEKTPPEPKEEVTIKANLI-----YADGKTQTA-----EFKGTFEATAEAY---RY 44
Db 1262 KEKLELEREKQLLRKKLLNCSIGNVKYSDFESDTSDSDSGDKGFKNPLYEPYNESEH 1321

QY 45 ADALKKDNGE---YTVDVADKGYTLN 67
Db 1322 TDILKKNNSPEYLYNDNVFORGECLN 1348

RESULT 11
Q27039 PRELIMINARY; PRT; 256 AA.
ID Q27039
AC Q27039
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Merozoite surface glycoprotein precursor (Fragment).
GN TPMS1.
OS Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Muguga;
RA Shields B.R., d'Oliviera C., McKellar S., Ben-Miled L., Kawazu S.,
RA Hide G.;
RT "Selection of Diversity at Putative Glycosylation sites in the
RT immunodominant merozoite/piroplasm surface antigen of Theileria
RT parasites.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z48740; CAA88632.1; -.
DR InterPro: IPR003407; Merozoite_Agen.
DR Pfam: PF02488; EMA; 1.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT NON_TER 256 256
SQ SEQUENCE 256 AA; 29528 MW; C982AEE71A45DC5B CRC64;

Query Match      18.7%; Score 69.5; DB 5; Length 256;
Best Local Similarity 31.7%; Pred. No. 40;
Matches 19; Conservative 10; Mismatches 30; Indels 1; Gaps 1;

QY 1 KEKTPPEPKEEVTIKANLIYADGKTQTAEFKGTFEATAEAYRYADALKKONGEYTVDA 60
Db 25 EEEKKKKKEDLTVDVTLSSWENVTSTPEAGGTLKAN-EGYRFKTKLVGDKTLYNVDT 83

```

[illegible]

```

DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 279 AA; 31379 MW; 07E356825E071E80 CRC64;

Query Match      18.7%; Score 69.5; DB 2; Length 279;
Best Local Similarity 30.0%; Pred. No. 44;
Matches 21; Conservative 9; Mismatches 17; Indels 23; Gaps 3;
Merozoite surface glycoprotein precursor (Fragment).
TPMS1.
GN Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Muguga;
RX MEDLINE=96089836; PubMed=8538686;
RA Shields B.R., d'Oliveira C., McKellar S., Ben-Miled L., Kawazu S.,
RA Hide G.;
RT "Selection of diversity at putative glycosylation sites in the
RT immunodominant merozoite/piroplasm surface antigen of Theileria
RT parasites.";
RL Mol. Biochem. Parasitol. 72:149-162(1995).
DR EMBL; U22889; AAC46910.1; -.
DR InterPro; IPR003407; Merozoite_Agen.
DR Pfam; PF02488; EMA; 1.
DR KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT NON_TER 256 256
SQ SEQUENCE 256 AA; 29528 MW; C982AEE71A45DC5B CRC64;

Query Match      18.7%; Score 69.5; DB 5; Length 256;
Best Local Similarity 31.7%; Pred. No. 40;
Matches 19; Conservative 10; Mismatches 30; Indels 1; Gaps 1;

QY 1 KEKTEPEEKVEITKANLIYADGKTQTAEFGTFTFEETAEAYRYADALKKDNGEYTDVA 60
DB 25 EEKKKKEDLTVDVTLSSWENVTSTPEAGGTLLKAN-EGYRFTLKVGDKTYLNVDTS 83

RESULT 13
Q9ADV8 PRELIMINARY; PRT; 279 AA.
AC Q9ADV8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Major outer membrane protein P30-II.
DN Ehrlichia canis.
OS Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=944;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oklahoma;
RX MEDLINE=98371112; PubMed=9705412;
RA Ohashi N., Unver A., Zhi N., Rikihisa Y.;
RT "Cloning and characterization of multigenes encoding the
RT immunodominant 30-kilodalton major outer membrane proteins of
RT Ehrlichia canis and application of the recombinant protein for
RT serodiagnosis.";
RL J. Clin. Microbiol. 36:2671-2680(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Okahoma;
RX MEDLINE=21153566; PubMed=11254561;
RA Ohashi N., Rikihisa Y., Unver A.;
RT "Analysis of transcriptionally Active Gene Clusters of Major Outer
RT Membrane Protein Multigene Family in Ehrlichia canis and E.
RT chaffeensis.";
RL Infect. Immun. 69:2083-2091(2001).
DR EMBL; AF078553; AAK28690.1; -.

```

